

QY 5 KPIADLYRGRESRPSAPR 22
:||||:|:|:|
Db 526 RPTADRRQGRDRRQADR 543

RESULT 8
T52297
squamosa promoter binding protein-homolog 5 [imported] - garden snapdragon (fragment)
C;Species: Antirrhinum majus (garden snapdragon)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52297
R;Cardon, G.H.; Hoehmann, S.; Klein, J.; Nettesheim, K.; Saedler, H.; Huijser, P.
Gene 237, 91-104, 1999
A;Title: Molecular characterisation of the Arabidopsis SBP-box genes.
A;Reference number: Z5236; MUID:99453765; PMID:10524240
A;Accession: T52297
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-323 <CAR>
A;Cross-references: UNIPROT:Q9SNV3; EMBL:AJ011623; PIDN:CAB56570.1
C;Genetics:
A;Gene: sbp5

Query Match 37.5%; Score 46.5; DB 2; Length 323;
Best Local Similarity 40.9%; Pred. No. 22;
Matches 9; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 1 PWQ---YKPIADLYRGRESRPS 19
|||:|:|:|
Db 170 PWQSDMQNVPFLRGTNRPS 191

RESULT 9
E70633
hypothetical protein Rv0398c - Mycobacterium tuberculosis (strain H37Rv)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: E70633
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70633
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-213 <COL>
A;Cross-references: UNIPROT:P95206; GB:Z84725; GB:AL123456; NID:G3261703; PIDN:CAB06589.
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv0398c
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0398c

Query Match 37.1%; Score 46; DB 2; Length 213;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 WOYKPIADLYRGRE 15
|||:|:|:|
Db 170 WYPPFFATRRGEE 193

RESULT 10
AH2650
transcription regulator, LysR family Atu0605 [imported] - Agrobacterium tumefaciens (strain C58)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AH2650
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, G.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.; Karp, P.; Romero, P.; Zhang, S.
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AH2650
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <KUR>
A;Cross-references: UNIPROT:Q8UHS4; GB:AE008688; PIDN:AA141622.1; PID:gl7738961; GSPDB: A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu0605
A;Map position: circular chromosome

Query Match 37.1%; Score 46; DB 2; Length 290;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 ADLYRGRESRPS 19
:||||:|:|:|
Db 279 SDLYAGRKSRPA 290

RESULT 11
G97432
probable transcription regulator (Pal138) [imported] - Agrobacterium tumefaciens (strain C58)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: G97432
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldmar, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: G97432
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-290 <KUR>
A;Cross-references: UNIPROT:Q8UHS4; GB:AE007869; PIDN:AAK86416.1; PID:gl5155552; GSPDB: C;Genetics:
A;Gene: AGR_C1075
A;Map position: circular chromosome

Query Match 37.1%; Score 46; DB 2; Length 290;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 ADLYRGRESRPS 19
:||||:|:|:|
Db 279 SDLYAGRKSRPA 290

RESULT 12
C71327
probable phosphoglucotomase - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: C71327
R;Praser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwi, rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McI they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: C71327
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-632 <COL>
A;Cross-references: UNIPROT:Q83428; GB:AE001219; GB:AE000520; NID:G3322693; PIDN:AA0654 A;Experimental source: strain Nichols
C;Genetics:

A:Gene: TP0413

Query Match 37.1%; Score 46; DB 2; Length 632;

Best Local Similarity 58.8%; Pred. No. 55;

Matches 10; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 4 YKPIADLYRGRESRPSA 20

Db 541 YRPI--LRGREQQDAA 555

RESULT 13

H84516

hypothetical protein At2g14410 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: H84516

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Mofatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487; PMID:10617197

A:Accession: H84516

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-679 <STO>

A:Cross-references: UNIPROT:Q9ZQ04; GB:AE002093; NID:G4263830; PIDN:AAD15473.1; GSPDB:GN

C:Genetics:

A:Gene: At2g14410

A:Map position: 2

Query Match 37.1%; Score 46; DB 2; Length 679;

Best Local Similarity 60.0%; Pred. No. 59;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 KPIADLYRGRESRPS 19

Db 289 EPIADLIRKRDPRS 303

RESULT 14

H83737

Glucosidase BH0704 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: H83737

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: AB3650; MUID:20512582; PMID:11058132

A:Accession: H83737

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-801 <STO>

A:Cross-references: UNIPROT:Q9K3Z5; GB:AP001509; GB:BA000004; NID:G10173176; PIDN:BA044

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH0704

C:Superfamily: lysosomal alpha-glucosidase; sucrase/isomaltase homology; trefoil homolog

Query Match 37.1%; Score 46; DB 2; Length 801;

Best Local Similarity 55.6%; Pred. No. 70;

Matches 10; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 1 PMQY---KPIADLYRGR 14

Db 567 PMQYPIVKRPSIDLMRFR 584

RESULT 15

JC2369

ribosomal protein L15, cytosolic [validated] - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004

C:Accession: JC2369; PC2233; JC2236; PC2158

R:Chan, Y.L.; Olivera, J.; Wool, I.G.

Biochem. Biophys. Res. Commun. 201, 108-114, 1994

A:Title: The primary structure of rat ribosomal protein L15.

A:Reference number: JC2236; MUID:94256965; PMID:8198562

A:Accession: JC2369

A:Molecule type: mRNA

A:Residues: 1-204 <CHA>

A:Cross-references: UNIPROT:P61314; EMBL:X78167; NID:G515864; PIDN:CAA55026.1; PID:G515

A:Accession: PC2233

A:Molecule type: protein

A:Residues: 2-26;137-173 <CH2>

A:Experimental source: liver

A:Note: the protein is designated as ribosomal protein L15

C:Genetics:

A:Gene: L15

C:Superfamily: rat ribosomal protein L15

C:Keywords: nucleotide binding; P-loop; phosphoprotein; protein biosynthesis; ribosome

F:2-204/Product: ribosomal protein L15 #status experimental <MAT>

F:126-131/Region: amyloid hexapeptide YKPFPE motif

F:163-171/Region: nucleotide-binding motif A (P-loop)

F:126/Binding site: phosphate (thr) (covalent) (by protein kinase C) #status predicted

Query Match

Best Local Similarity 36.7%; Score 45.5; DB 1; Length 204;

Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY 2 WQYKPIADLYRG-RESRPSAPR 22

Db 28 WQYQLSALHRAPEPTRPDKAR 49

Search completed: November 10, 2004, 14:52:16

Job time : 5.39623 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 29.3057 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-47

Perfect score: 124

Sequence: 1 PMQYKPTADLYRGSRPSAPR 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trnml.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	43.5	187	Q9YE21	Q9YE21 aeropyrum p
2	52	41.9	211	Q6H016	Q6H016 fremyella d
3	52	41.9	394	Q7ZXX1	Q7ZXX1 xenopus lae
4	51	41.1	458	Q2N883	Q2N883 plasmodium
5	51	41.1	556	1 BETA_XANAC	Q8PPG8 xanthomonas
6	51	41.1	1017	2 Q8VW43	Q8VW43 bradyrhizob
7	51	41.1	1483	2 Q6L420	Q6L420 solanum dem
8	51	41.1	1483	2 AAT39963	AAT39963 solanum d
9	50.5	40.7	163	2 Q6TJN2	Q6TJN2 bovine papu
10	50.5	40.7	488	2 P72766	P72766 synechocyst
11	50.5	40.7	683	2 Q6TVAS	Q6TVAS bovine papu
12	50.5	40.7	683	2 AAR98436	AAR98436 bovine pa
13	50	40.3	197	2 Q7Q7A9	Q7Q7A9 anopheles g
14	50	40.3	217	2 Q9ZVH6	Q9ZVH6 arabidopsis
15	50	40.3	282	2 Q9M0T7	Q9M0T7 arabidopsis
16	50	40.3	441	2 Q82BH1	Q82BH1 streptomyce
17	50	40.3	637	2 Q8H8M4	Q8H8M4 oryza sativ
18	50	40.3	747	2 Q8S9Y3	Q8S9Y3 arabidopsis
19	49	39.5	152	2 Q9CY74	Q9CY74 mus musculu
20	49	39.5	348	2 Q8N5E2	Q8N5E2 homo sapien
21	49	39.5	348	2 Q7TPP7	Q7TPP7 mus musculu
22	49	39.5	364	2 Q6ZUE3	Q6ZUE3 homo sapien
23	49	39.5	364	2 BAC86282	BAC86282 homo sapi
24	49	39.5	556	1 BETA_XANCP	Q8P5D7 xanthomonas
25	49	39.5	560	2 Q6LW44	Q6LW44 homo sapien
26	49	39.5	560	2 AAH64600	AAH64600 homo sapi
27	49	39.5	1116	2 Q8CHD2	Q8CHD2 mus musculu
28	49	39.5	1163	2 Q94872	Q94872 homo sapien
29	49	39.5	1337	2 Q8A3U4	Q8A3U4 bacteroides
30	48.5	39.1	136	2 Q7T3P3	Q7T3P3 acipenser s
31	48.5	39.1	203	1 RL15_WHITE	P30736 chironomus

32 48.5 39.1 203 1 RL15_EPICO P61367 epinephelus
33 48.5 39.1 203 1 RL15_MONAL Q7L3N2 monopterus
34 48.5 39.1 203 1 RL15_SINKN Q7L3M9 siniperca k
35 48.5 39.1 204 2 Q6T278 Q6T278 acipenser g
36 48.5 39.1 204 2 Q6T279 Q6T279 acipenser s
37 48.5 39.1 204 2 Q6T280 Q6T280 acipenser s
38 48.5 39.1 204 2 AAS59857 AAS59857 acipenser
39 48.5 39.1 204 2 AAS59858 AAS59858 acipenser
40 48.5 39.1 204 2 AAS59859 AAS59859 acipenser
41 48.5 39.1 281 2 O8GLM6 O8GLM6 anaplasma o
42 48 38.7 420 1 INX3_CAEEL Q19746 caenorhabdi
43 48 38.7 833 2 Q7RW59 Q7RW59 neurospora
44 48 38.7 1225 2 Q7S641 Q7S641 neurospora
45 48 38.7 1225 2 CAE85584 CAE85584 neurospor

ALIGNMENTS

RESULT 1

Q9YE21 PRELIMINARY; PRT; 187 AA.
ID Q9YE21
AC Q9YE21;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APO0749.
GN OrderedLocusNames=AP0749;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcates;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA79726.1; -.
DR FIR; F72665; F72665.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 187 AA; 20537 MW; 8C179729470E2A9 CRC64;

Query Match 43.5%; Score 54; DB 2; Length 187;
Best Local Similarity / 57.9%; Pred. No. 5.9;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 YKPIADLYRGSRPSAPR 22

DB 35 YKLVGLLGLRLPKLPR 53

RESULT 2

Q6H016 PRELIMINARY; PRT; 211 AA.
ID Q6H016
AC Q6H016;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE FraH-like protein.
OS Fremyella diplosiphon (Calothrix PCC 7601).
OC Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Fremyella.
OX NCBI_TaxID=1197;
RN [1]
RP SEQUENCE FROM N.A.
SQ STRAIN=FD33;

[3] SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC044084; AAH44084.1; -
DR KSSP; P13590; IIES
DR InterPro; IPRO07110; IG-like.
DR InterPro; IPRO03598; IG_C2.
DR InterPro; IPRO03585; Neurexin-like.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00294; 4.1m; 1.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS0835; IG_LIKE; 3.
SQ SEQUENCE 394 AA; 42730 MW; F141D8E6869254A CRC64;

Query Match 41.9%; Score 52; DB 2; Length 394;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 12; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 3 QYKPIADLYRGESRPSAPR 22
||| | : |||||
Db 225 QYKPTAKI----ESRPSMPR 240

RESULT 4
Q9N883 PRELIMINARY; PRT; 458 AA.

ID Q9N883
AC Q9N883
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Vir28 protein.
GN Namevir28;
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP OLIVER K., Bowman S., Hall N., Quail M., Rajandream M.A., Harris D., del Portillo H.A., Lencer M., Barrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL360354; CAB96715.1; -
SQ SEQUENCE 458 AA; 53134 MW; 785060C82705BB16 CRC64;

Query Match 41.1%; Score 51; DB 2; Length 458;
Best Local Similarity 40.0%; Pred. No. 45;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 WQYKPIADLYRGESRPSAP 21
::| : |||||
Db 407 FRYPVGTFFRGRRPGIP 426

RESULT 5
BETA_XANAC STANDARD; PRT; 556 AA.

ID BETA_XANAC
AC Q8PPG8;
DT 29-MAR-2004 (Rel. 43; Created)
DT 29-MAR-2004 (Rel. 43; Last sequence update)
DT 05-JUL-2004 (Rel. 44; Last annotation update)
DE Choline dehydrogenase (EC 1.1.99.1) (CHD).
GN Name-beta; Ordered locus Names=XAC0718;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=250224215; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R., Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,


```

RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Chapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.P.,
RA Spinola L.A.F., Takita M.A., Tamura R.B., Teixeira B.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Tubal J.C., Kitajima J.P.,
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
CC -1- FUNCTION: Can catalyze the oxidation of choline to betaine
CC aldehyde and betaine aldehyde to glycine betaine (by similarity).
CC -1- CATALYTIC ACTIVITY: Choline + acceptor = betaine aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: FAD (by similarity).
CC -1- PATHWAY: Betaine biosynthesis from choline; first step.
CC -1- SIMILARITY: Belongs to the GMC oxidoreductase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE011702; AAM35607.1; -.
DR HAMAP; MF 00750; -.
DR InterPro; IPR000172; GMC_oxred.
DR InterPro; IPR007867; GMC_oxred_C.
DR Pfam; PF05199; GMC_oxred_C; 1.
DR Pfam; PF00732; GMC_oxred_N; 1.
DR TIGRFAMs; TIGR01810; betaA; 1.
DR PROSITE; PS00623; GMC_OXRED_1; 1.
DR PROSITE; PS00624; GMC_OXRED_2; 1.
KW Complete proteome; FAD; Flavoprotein; Oxidoreductase.
FT NP_BIND 6 35 FAD (ADP part) (Probable).
FT ACT_SITE 475 475 By similarity.
SQ SEQUENCE 556 AA; 61265 MW; 026B7F3EE59B8DA CRC64;

Query Match 41.1%; Score 51; DB 1; Length 556;
Best Local Similarity 62.5%; Pred. No. 55;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Cy 5 KPIADLYRGRESRPSA 20
Db 439 QPALDAYRGRESRPSA 454

RESULT 6
ID Q8VW43 PRELIMINARY; PRT; 1017 AA.
AC Q8VW43;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Proline dehydrogenase.
GN Name=PutA;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GX201;
RA Wu B., Tang X., Bai X., Tang D., Lu A., Tang J., Ma Q.;
RA Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

Query Match 41.1%; Score 51; DB 1; Length 1483;
Best Local Similarity 55.6%; Pred. No. 11e+02;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 5 KPIADLYRGRESRPSA 22
Db 526 RPTADRRQGRSRRPQHR 543

RESULT 7
ID Q6L420 PRELIMINARY; PRT; 1483 AA.
AC Q6L420;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Putative polyprotein.
GN Name=PGEC989P08.7;
OS Solanum demissum (Wild potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=50514;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell R.C., Liu J., Zaborsky J., Tallon L., Baker B.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ronning C.M.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC146506; AAT39963.1; -.
DR InterPro; IPR000923; BlueCu_1.
DR InterPro; IPR009007; Pept_Aspartic.
DR InterPro; IPR001584; Rve.
DR Pfam; PF00655; rve; 1.
DR Pfam; PF00078; RVT; 1.
DR PROSITE; PS00196; COPPER BLUE; UNKNOWN_1.
KW Polyprotein; RNA-directed DNA polymerase; Transferase.
SQ SEQUENCE 1483 AA; 169333 MW; 30D527FF57244348 CRC64;

```

Best Local Similarity 46.7%; Pred. No. 1.6e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 PWQYKPIADLYRGRE 15
|||:|:|:|:|:|:|
D5 75 PWKYPTVTYKGE 89

RESULT 8
AAT39963 PRELIMINARY; PRT; 1483 AA.
AC AAT39963;
DT 01-JUN-2004 (TREMELrel. 27, Created)
DT 01-JUN-2004 (TREMELrel. 27, Last sequence update)
DT 01-JUN-2004 (TREMELrel. 27, Last annotation update)
DE Putative polyprotein.
GN PGEC989P08.7.
OS Solanum demissum (Wild potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lauids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=50514;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell R.C., Liu J., Zaborsky J., Tallon L., Baker B.;
RT "Solanum demissum chromosome 5 BAC PGEC989P08 genomic sequence.";
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Rønning C.M.;
RT Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Buell R.;
RT Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC146506; AAT39963.1; -.
KW Polyprotein.
SQ SEQUENCE 1483 AA; 169333 MW; 30D527FF57244348 CRC64;

Query Match 41.1%; Score 51; DB 2; Length 1483;
Best Local Similarity 46.7%; Pred. No. 1.6e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 PWQYKPIADLYRGRE 15
|||:|:|:|:|:|:|
D5 75 PWKYPTVTYKGE 89

RESULT 9
Q6JIN2 PRELIMINARY; PRT; 163 AA.
AC Q6JIN2;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE Putative viral core protein P4b (fragment).
OS Bovine papular stomatitis virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Parapoxvirus.
OX NCBI_TaxID=129727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V660;
RX PubMed=15166423;
RA Tikkanen M.K., McInnes C.J., Mercer A.A., Buttner M., Tuimala J.,
RA Hirvelä-Koski V., Neuvonen E., Huovilainen A.;
RT "Recent isolates of parapoxvirus of Finnish reindeer (Rangifer
tarandus tarandus) are closely related to bovine pseudocowpox virus.";
RL J. Gen. Virol. 85:1413-1418(2004).

EMBL; AY453689; AAS13365.1; -.
FT NON-TER 1
SQ SEQUENCE 163 AA; 17939 MW; 1630AB5779679861 CRC64;

Query Match 40.7%; Score 50.5; DB 2; Length 163;
Best Local Similarity 61.1%; Pred. No. 17;
Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 PWQYKPIADLYRG-RESR 17
|||:|:|:|:|:|:|
D5 57 PWLYDPMGALSARGAESR 74

RESULT 10
P72766 PRELIMINARY; PRT; 488 AA.
AC P72766;
DT 01-FEB-1997 (TREMELrel. 02, Created)
DT 01-FEB-1997 (TREMELrel. 02, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE S111686 protein.
GN OrderedLocustNames=s111686;
OS Synchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
Synchocystis sp. strain PCC6803. II. Sequence determination of the
entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
DR EMBL; D90900; BAA16781.1; -.
DR PIR; S74629; S74629.
KW Complete proteome.
SQ SEQUENCE 488 AA; 57146 MW; 578A0B158BE99766 CRC64;

Query Match 40.7%; Score 50.5; DB 2; Length 488;
Best Local Similarity 44.4%; Pred. No. 57;
Matches 8; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 1 PWQYKPIADLYRGRESR 18
|||:|:|:|:|:|:|
D5 130 PWQYRPVS-IQLGRPKP 146

RESULT 11
Q6TVA9 PRELIMINARY; PRT; 683 AA.
AC Q6TVA9;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE ORF079 virion core protein P4b.
OS Bovine papular stomatitis virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Parapoxvirus.
OX NCBI_TaxID=129727;

```

RN  RC STRAIN=PEST;
RP  RA Anopheles Genome Sequencing Consortium;
RC  RA Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
RX  CC -!- CAUTION: The sequence shown here is derived from an
RA  EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
PA  preliminary data.
RA  DR EMBL; AAAB01008960; EAA1143.1; -.
SQ  SEQUENCE 197 AA; 22589 MW; BB877FB775441505 CRC64;

Query Match 40.3%; Score 50; DB 2; Length 197;
Best Local Similarity 61.9%; Pred. No. 25;
Matches 13; Conservative 2; Mismatches 2; Indels 4; Gaps 2;

QY 5 KPI-ADLYRGRES---RFSAP 21
Db 7 KPVRVRYQGRESIFKRPVAP 27

RESULT 14
Q9ZVH6 PRELIMINARY; PRT; 217 AA.
ID Q9ZVH6
AC Q9ZVH6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative retroelement pol polyprotein.
GN Name=At2g38520;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AC005499; AAC67366.1; -.
DR PIR; A84806; A84806.
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
KW Polyprotein.
SQ SEQUENCE 217 AA; 24148 MW; 7B5D12B178D57EFD CRC64;

Query Match 40.3%; Score 50; DB 2; Length 217;
Best Local Similarity 47.4%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 PWQKPIADLYRGRESRPS 19
Db 178 PEDYKPIVDQMEGRDTPN 196

RESULT 15
Q9MOT7 PRELIMINARY; PRT; 282 AA.
ID Q9MOT7
AC Q9MOT7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein At4g07710.
GN Name=At4g07710;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RN  RC STRAIN=PEST;
RP  RA Anopheles Genome Sequencing Consortium;
RC  RA Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
RX  CC -!- CAUTION: The sequence shown here is derived from an
RA  EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is
PA  preliminary data.
RA  DR EMBL; AAAB01008960; EAA1143.1; -.
SQ  SEQUENCE 197 AA; 22589 MW; BB877FB775441505 CRC64;

Query Match 40.7%; Score 50.5; DB 2; Length 683;
Best Local Similarity 61.1%; Pred. No. 82;
Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 PWQKPIADLYRG-RRSR 17
Db 577 PWLYDPMALSARGARESR 594

RESULT 12
AAR98436 PRELIMINARY; PRT; 683 AA.
ID AAR98436
AC AAR98436;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE ORF079 virion core protein p4b.
OS Bovine papular stomatitis virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Parapoxvirus.
OX NCBI_TaxID=129727;
RN [1]
RP SEQUENCE FROM N.A.
RA Delhon G., Tulman E.R., Afonso C.L., Lu Z., Piccone M.E., Kutish G.F.,
RA de la Concha-Bermejillo A., Lehmkuhl H.D.;
RX PubMed=14671098;
RA Rock D.L.;
RL "Genomes of the Parapoxviruses Orf Virus and Bovine Papular
RT Stomatitis Virus.";
RL J. Virol. 78:168-177(2004).
DR EMBL; AY386265; AAR98436.1; -.
DR InterPro; IPR004972; Pox_P4B.
DR Pfam; PF03292; Pox_P4B; 1.
DR ACPC6267.
GN Name=agCG50481; ORFNames=ENSANGC00000011638;
OS Anopharies Gambae str. PSF1.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
```

RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMEL; AL161507; CAB77930.1; -
 DR FIR; A85076; A85076.
 KW Hypothetical protein.
 SQ SEQUENCE 282 AA; 31518 MW; 1A691C1C95991A03 CRC64;

Query Match 40.3%; Score 50; DB 2; Length 282;
 Best Local Similarity 60.0%; Pred. No. 37;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 KPIADLYRGRESRPS 19
 :|||||:|:|
 Db 188 EPIADLFRKKRERS 202

Search completed: November 10, 2004, 14:50:04
 Job time : 31.3057 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 6.58868 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-48

Perfect score: 91

Sequence: 1 LFSVLLRYLADNPLPGGS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	52.7	1167	1	US-08-100-709-2
2	48	52.7	1167	1	US-08-176-855-2
3	48	52.7	1167	1	US-08-474-038-2
4	48	52.7	1167	2	US-08-779-046-2
5	48	52.7	1167	2	US-08-881-340-2
6	48	52.7	1168	1	US-08-448-170-6
7	48	52.7	1168	3	US-08-961-803-8
8	48	52.7	1168	4	US-09-661-322A-22
9	48	52.7	1170	1	US-08-032-364-2
10	46	50.5	696	4	US-09-489-039A-12404
11	45	49.5	630	3	US-08-771-986A-2
12	45	49.5	630	3	US-08-769-802A-2
13	45	49.5	773	4	US-09-328-352-8203
14	45	49.5	823	3	US-08-461-551-2
15	45	49.5	823	3	US-09-037-621A-2
16	45	49.5	823	4	US-09-583-717-2
17	45	49.5	1163	1	US-08-239-474A-11
18	45	49.5	1163	2	US-08-732-495-11
19	45	49.5	1163	3	US-09-178-252-21
20	45	49.5	1163	4	US-09-826-660-21
21	45	49.5	1189	1	US-07-828-788A-16
22	45	49.5	1189	1	US-08-356-034-6
23	45	49.5	1189	1	US-08-602-737-2
24	45	49.5	1189	2	US-08-980-071-2
25	45	49.5	1189	2	US-08-980-071-4
26	45	49.5	1189	2	US-08-980-071-6
27	45	49.5	1189	2	US-08-980-071-8

28	45	49.5	1189	2	US-08-980-071-10	Sequence 10, Appl
29	45	49.5	1189	2	US-08-980-071-12	Sequence 12, Appl
30	45	49.5	1189	2	US-08-980-071-59	Sequence 59, Appl
31	45	49.5	1189	2	US-08-980-071-61	Sequence 61, Appl
32	45	49.5	1189	2	US-08-757-536-2	Sequence 2, Appl
33	45	49.5	1189	2	US-08-757-536-4	Sequence 4, Appl
34	45	49.5	1189	2	US-08-757-536-6	Sequence 6, Appl
35	45	49.5	1189	2	US-08-757-536-8	Sequence 8, Appl
36	45	49.5	1189	2	US-08-757-536-10	Sequence 10, Appl
37	45	49.5	1189	3	US-08-757-536-12	Sequence 12, Appl
38	45	49.5	1189	3	US-09-314-093-2	Sequence 2, Appl
39	45	49.5	1189	3	US-09-314-093-4	Sequence 4, Appl
40	45	49.5	1189	3	US-09-314-093-6	Sequence 6, Appl
41	45	49.5	1189	3	US-09-314-093-8	Sequence 8, Appl
42	45	49.5	1189	3	US-09-314-093-10	Sequence 10, Appl
43	45	49.5	1189	3	US-09-314-093-12	Sequence 12, Appl
44	45	49.5	1189	3	US-09-314-093-59	Sequence 59, Appl
45	45	49.5	1189	3	US-09-314-093-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1

US-08-100-709-2
; Sequence 2, Application US/08100709
; Patent No. 5322687

GENERAL INFORMATION:

APPLICANT: Donovan, William P.
APPLICANT: Tan, Yiping
APPLICANT: Jany, Christine S.
APPLICANT: Gonzalez Jr., Jose M.
TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET4 AND CRYET5
TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
ADDRESS: Nadel
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100,709
FILING DATE: 19930729
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Egolf, Christopher
REGISTRATION NUMBER: 27633
REFERENCE/DOCKET NUMBER: 7205-49
TELEPHONE: 215-757-1590
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1167 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-100-709-2

Query Match 52.7%; Score 48; DB 1; Length 1167;

Best Local Similarity 61.5%; Pred. No. 20; Mismatches 4; Indels 0; Gaps 0;

Matches 8; Conservative 4;

QY 5 LRLRYLADNPLPGG 17

Db 41 LRLRYLADNPLPGG 53

```

RESULT 2
US-08-176-865-2
; Sequence 2, Application US/08176865
; Patent No. 5616319
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yuping
; APPLICANT: Jany, Christine S.
; APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESSEE: Nadel
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,865
; FILING DATE: 30-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/100,709
; FILING DATE: 29-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Egolf, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1167 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-176-865-2

Query Match 52.7%; Score 48; DB 1; Length 1167;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LLRYLADNPLPGG 17
Db 41 LLQFLNNFVPGG 53

RESULT 3
US-08-474-038-2
; Sequence 2, Application US/08474038
; Patent No. 5679343
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yuping
; APPLICANT: Jany, Christine S.
; APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESSEE: Nadel
; STREET: 1601 Market Street, 36th Floor

```

```

; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,038
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER: US 08/176,865
; FILING DATE: 30-DEC-1993
; APPLICATION NUMBER: US 08/100,709
; FILING DATE: 29-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Egolf, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1167 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-474-038-2

Query Match 52.7%; Score 48; DB 1; Length 1167;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LLRYLADNPLPGG 17
Db 41 LLQFLNNFVPGG 53

RESULT 4
US-08-779-046-2
; Sequence 2, Application US/08779046
; Patent No. 5854053
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yuping
; APPLICANT: Jany, Christine S.
; APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESSEE: Nadel
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,046
; FILING DATE: 06-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 08/100,709
; FILING DATE: 29-JUL-1993

```

ATTORNEY/AGENT INFORMATION:
 NAME: Egolf, Christopher
 REGISTRATION NUMBER: 27633
 REFERENCE/DOCKET NUMBER: 7205-49
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-757-1590
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1167 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-779-046-2

Query Match 52.7%; Score 48; DB 2; Length 1167;
 Best Local Similarity 61.5%; Pred. No. 20;
 Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
 DB 41 LLQFLNNFVPGG 53

RESULT 5

US-08-881-340-2
 Sequence 2, Application US/08881340
 Patent No. 5942658
 GENERAL INFORMATION:
 APPLICANT: Donovan, William P.
 APPLICANT: Tan, Yiping
 APPLICANT: Jany, Christine S.
 APPLICANT: Gonzalez Jr., Jose M.
 TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
 TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
 ADDRESSEE: Nadel
 STREET: 1601 Market Street, 36th Floor
 CITY: Philadelphia
 STATE: Pennsylvania
 COUNTRY: U.S.A.
 ZIP: 19103

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/881,340
 FILING DATE: 24-JUN-1997
 CLASSIFICATION: 424

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/100,709
 FILING DATE: 29-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Egolf, Christopher
 REGISTRATION NUMBER: 27633
 REFERENCE/DOCKET NUMBER: 7205-49
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-757-1590
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1167 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-881-340-2

Query Match 52.7%; Score 48; DB 2; Length 1167;
 Best Local Similarity 61.5%; Pred. No. 20;
 Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
 DB 41 LLQFLNNFVPGG 53

RESULT 6

US-08-448-170-6
 Sequence 6, Application US/08448170
 Patent No. 5723758
 GENERAL INFORMATION:
 APPLICANT: Payne, Jewel
 APPLICANT: Cummings, David A.
 APPLICANT: Cannon, Raymond J.C.
 APPLICANT: Narva, Kenneth E.
 APPLICANT: Stelman, Steve
 TITLE OF INVENTION: No. 5723758el Bacillus thuringiensis Isolate Denoted
 TITLE OF INVENTION: B.t. PSI58C2, Active Against Lepidopteran Pests, and Genes
 TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: David R. Saliwanchik
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: Florida
 COUNTRY: USA
 ZIP: 32606

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/448,170
 FILING DATE:
 CLASSIFICATION: 424

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/069,902
 FILING DATE: 01-JUNE-1993
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/759,247
 FILING DATE: 13-SEPT-1991
 CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
 NAME: Saliwanchik, David R.
 REGISTRATION NUMBER: 31,794
 REFERENCE/DOCKET NUMBER: M/S 102D.C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (904) 375-8100
 TELEFAX: (904) 372-5800
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1168 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-448-170-6

Query Match 52.7%; Score 48; DB 1; Length 1168;
 Best Local Similarity 61.5%; Pred. No. 20;
 Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
 DB 41 LLQFLNNFVPGG 53

RESULT 7

US-08-961-803-8
 Sequence 8, Application US/08961803
 Patent No. 6150589
 GENERAL INFORMATION:

```
; APPLICANT: Payne, Jewel
; APPLICANT: Cummings, David A.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stelman, Steve
; TITLE OF INVENTION: No. 6150589el Bacillus thuringiensis Isolate Denoted
; TITLE OF INVENTION: B.t. PS158C2, Active Against Lepidopteran Pests, and Genes
; TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jay M. Sanders
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,803
; FILING DATE: 31-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/069,902
; FILING DATE: 01-JUNE-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,247
; FILING DATE: 13-SEPT-1991
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/448,170
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: M/S 102DCD1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1168 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-961-803-8

Query Match 52.7%; Score 48; DB 3; Length 1168;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
Db 41 LLQFLNNFVPGG 53

RESULT 8
US-09-661-322A-22
; Sequence 22, Application US/09661322A
; Patent No. 6593293
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rupa, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin Compos
```

```
; TITLE OF INVENTION: and Methods of Use
; FILE REFERENCE: MECO201
; CURRENT APPLICATION NUMBER: US/09/661,322A
; CURRENT FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; US-09-661-322A-22

Query Match 52.7%; Score 48; DB 4; Length 1168;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
Db 41 LLQFLNNFVPGG 53

RESULT 9
US-08-032-364-2
; Sequence 2, Application US/08032364
; Patent No. 5356623
; GENERAL INFORMATION:
; APPLICANT: von Tersch, Michael A.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS CRYET1 TOXIN GENE
; TITLE OF INVENTION: AND PROTEIN TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESSEE: Nagel
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/032,364
; FILING DATE: 19930317
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Egolf, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-31
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1170 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-032-364-2

Query Match 52.7%; Score 48; DB 1; Length 1170;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
Db 41 LLQFLNNFVPGG 53

RESULT 10
US-09-489-039A-12404
; Sequence 12404, Application US/09489039A
```


; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12404
; LENGTH: 696
; TYPE: PRT
; ORGANISM: klebsiella pneumoniae
US-09-489-039A-12404

Query Match 50.5%; Score 46; DB 4; Length 696;
Best Local Similarity 58.8%; Pred. No. 24;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 FSVLLRYLADNPLPGS 18
Db 237 FVSLRISADDFLRKGN 253

RESULT 11
US-08-771-986A-2
; Sequence 2, Application US/08771986A
; Patent No. 6043415
; GENERAL INFORMATION:
; APPLICANT: STRIZHOV, Nicolai
; APPLICANT: SCHELL, Jeff
; APPLICANT: ZILBERTSTEIN, Aviah
; APPLICANT: KELLER, Menachem
; APPLICANT: SNEH, Baruch
; APPLICANT: KONCZ, Csaba
; TITLE OF INVENTION: SYNTHETIC BACILLUS THURINGIENSIS GENE
; TITLE OF INVENTION: ENCODING AN INSECT TOXIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/771,986A
; FILING DATE: 23-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: FEEL, Richard C.
; REGISTRATION NUMBER: 35,792
; REFERENCE/DOCKET NUMBER: 026433/0118
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-771-986A-2

Query Match 49.5%; Score 45; DB 3; Length 630;
Best Local Similarity 53.8%; Pred. No. 31;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Qy 5 LLRYLADNPLPGG 17
Db 41 LVQFLVSNFVPGG 53
RESULT 12
US-08-769-802A-2
; Sequence 2, Application US/08769802A
; Patent No. 6110668
; GENERAL INFORMATION:
; APPLICANT: Strizhov, Nicolai; Koncz, Csaba; Schell, Jeff
; TITLE OF INVENTION: Gene Synthesis Method
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wigman, Cohen, Leitner & Myers, P.C.
; STREET: 900 17th Street, N.W., Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,802A
; FILING DATE: 20 - DEC - 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,896
; FILING DATE: 07 - OCT - 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Cohen, Herbert
; REGISTRATION NUMBER: 25,109
; REFERENCE/DOCKET NUMBER: 0363.004/P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 463 - 7700
; TELEFAX: (202) 463 - 6915
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 amino acid residues
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: truncated protein
; DESCRIPTION: truncated protein
; ANTI-SENSE: no
; FRAGMENT TYPE: N-terminal fragment
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: K26-21, Mri-37, subsp. aizawai 7.29
; INDIVIDUAL ISOLATE: not applicable
; HAPLOTYPE: not applicable
; TISSUE TYPE: not applicable
; CELL TYPE: unicellular organism
; IMMEDIATE SOURCE: strains K26-21, Mri-37, subsp. aizawai
; IMMEDIATE SOURCE: 7.29
; CLONE: not applicable
; LIBRARY: not applicable
; POSITION IN GENOME: not applicable
; CHROMOSOME/SEGMENT: not applicable
; MAP POSITION: not applicable
; UNITS: not applicable
; FEATURE: insecticidal control protein
; NAME/KEY: SEQ. ID NO. 2: is the sequence in claim 6, and
; NAME/KEY: relevant residues from 1 to 630
; LOCATION: not available
; IDENTIFICATION METHOD: by experiment
; OTHER INFORMATION: specifically toxic to insects of spodoptera
; OTHER INFORMATION: genus confers Spodoptera resistance being expressed in trans

```

1 COMPUTER READABLE FORM:
2 MEDIUM TYPE: Floppy disk
3 COMPUTER: IBM PC compatible
4 OPERATING SYSTEM: PC-DOS/MS-DOS
5 SOFTWARE: Patent In Release #1.0, Version #1.25
6 CURRENT APPLICATION DATA:
7 APPLICATION NUMBER: US/08/461,551
8 FILING DATE: 05-JUN-1995
9 CLASSIFICATION: 435
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: US 08/251,652
12 FILING DATE: 31-MAY-1994
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: US 07/458,754
15 FILING DATE: 11-DEC-1989
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: EP 88 401 121.4
18 FILING DATE: 06-MAY-1988
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: FR 87 08090
21 FILING DATE: 10-JUN-1987
22 ATTORNEY/AGENT INFORMATION:
23 NAME: HUNTINGTON, R. D.
24 REGISTRATION NUMBER: 27,503
25 REFERENCE/DOCKET NUMBER: 010830-073
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (703) 836-6620
28 TELEFAX: (703) 836-2021
29 INFORMATION FOR SEQ ID NO: 2:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 823 amino acids
32 TYPE: amino acid
33 STRANDEDNESS: unknown
34 TOPOLOGY: unknown
35 MOLECULE TYPE: peptide
36 US-08-461-551-2
37
38 Query Match 49.5%; Score 45; DB 1; Length 823;
39 Best Local Similarity 53.8%; Pred. No. 42;
40 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
41
42 QY 5 LLRYLADNPLPGG 17
43 DB 41 LVQFLVSNFVPGG 53
44
45 RESULT 15
46 US-09-037-621A-2
47 ; Sequence 2, Application US/09037621A
48 ; Patent No. 6310035
49 ; GENERAL INFORMATION:
50 ; APPLICANT: SANCHIS, Vincent
51 ; LERECCLUS, Didier
52 ; MENOUD, Ghislaine
53 ; LECADET, Marguerite-Marie
54 ; MARTOURET, Daniel
55 ; DEDONDER, Raymond
56 ;
57 TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR
58 ; A LARVICIDAL ACTIVITY TOWARDS
59 ; LEPIDOPTERA
60
61 NUMBER OF SEQUENCES: 2
62 CORRESPONDENCE ADDRESS:
63 ADDRESSES: BURNS, DOANE, SWECKER & MATHIS
64 STREET: P.O. Box 1404
65 CITY: Alexandria
66 STATE: Virginia
67 COUNTRY: USA
68 ZIP: 22313
69
70 COMPUTER READABLE FORM:
71 MEDIUM TYPE: Floppy disk
72 COMPUTER: IBM PC compatible
73 OPERATING SYSTEM: PC-DOS/MS-DOS
74 SOFTWARE: Patent In Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,621A
FILING DATE: 10-Mar-1998
CLASSIFICATION: <Unknown>
11-DEC-1989
06-MAY-1988
10-JUN-1987

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/461,551
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/458,754
FILING DATE: 11-DEC-1989
APPLICATION NUMBER: EP 88 401 121.4
FILING DATE: 06-MAY-1988
APPLICATION NUMBER: FR 87 08090
FILING DATE: 10-JUN-1987

ATTORNEY/AGENT INFORMATION:
NAME: HUNTINGTON, R. D.
REGISTRATION NUMBER: 27,903
REFERENCE/POCKET NUMBER: 010830-073
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 823 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-037-621A-2

```

```

Query Match      49.5%; Score 45; DB 3; Length 823;
Best Local Similarity 53.8%; Pred. No. 42;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      5 LLRYLADNPLPG 17
Db      41 LVQFLVSNFVPG 53

```

Search completed: November 10, 2004, 14:55:36
Job time.: 7.63868 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 20.8189 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-48

Perfect score: 91
Sequence: 1 LFSVLLRYLADNPLPGGS 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	100.0	18	14	US-10-092-750-48
2	48	52.7	1168	14	US-10-428-961-22
3	48	52.7	1168	15	US-10-614-524-4
4	46	50.5	97	17	US-10-425-115-281976
5	46	50.5	318	16	US-10-437-963-173457
6	46	50.5	619	17	US-10-425-115-259841
7	46	50.5	622	15	US-10-425-114-37487
8	46	50.5	622	15	US-10-425-114-60096
9	46	50.5	622	15	US-10-425-114-72803
10	45	49.5	71	15	US-10-424-599-147749
11	45	49.5	823	10	US-09-918-485-2
12	45	49.5	1163	9	US-09-826-660-21
13	45	49.5	1189	10	US-09-972-175-2

14	45	49.5	1189	10	US-09-972-175-4	Sequence 4, Appli
15	45	49.5	1189	10	US-09-972-175-6	Sequence 6, Appli
16	45	49.5	1189	10	US-09-972-175-8	Sequence 8, Appli
17	45	49.5	1189	10	US-09-972-175-10	Sequence 10, Appli
18	45	49.5	1189	10	US-09-972-175-12	Sequence 12, Appli
19	45	49.5	1189	10	US-09-972-175-59	Sequence 59, Appli
20	45	49.5	1189	10	US-09-972-175-61	Sequence 61, Appli
21	45	49.5	1189	11	US-09-837-961-6	Sequence 20, Appli
22	45	49.5	1189	14	US-10-102-469-20	Sequence 6, Appli
23	45	49.5	1189	14	US-10-200-532-2	Sequence 2, Appli
24	45	49.5	1189	14	US-10-200-532-4	Sequence 4, Appli
25	45	49.5	1189	14	US-10-200-532-6	Sequence 6, Appli
26	45	49.5	1189	14	US-10-200-532-8	Sequence 8, Appli
27	45	49.5	1189	14	US-10-200-532-10	Sequence 10, Appli
28	45	49.5	1189	14	US-10-200-532-12	Sequence 12, Appli
29	45	49.5	1189	14	US-10-200-532-59	Sequence 59, Appli
30	45	49.5	1189	14	US-10-200-532-61	Sequence 6, Appli
31	45	49.5	1189	17	US-10-825-751-6	Sequence 7, Appli
32	45	49.5	1189	17	US-10-782-030-7	Sequence 9, Appli
33	45	49.5	1189	17	US-10-782-036-9	Sequence 40847, A
34	44	48.4	164	16	US-10-767-701-40547	Sequence 14, Appli
35	44	48.4	352	9	US-09-073-009-14	Sequence 14, Appli
36	44	48.4	352	9	US-09-023-588-14	Sequence 14, Appli
37	44	48.4	352	9	US-09-793-306-14	Sequence 14, Appli
38	43	47.3	62	15	US-10-424-599-144449	Sequence 144449,
39	43	47.3	89	17	US-10-425-115-218707	Sequence 218707,
40	43	47.3	225	17	US-10-425-115-229739	Sequence 229739,
41	43	47.3	248	15	US-10-425-114-64833	Sequence 64833, A
42	43	47.3	268	9	US-09-841-132-432	Sequence 432, App
43	43	47.3	355	9	US-09-922-683-10	Sequence 10, Appli
44	43	47.3	686	16	US-10-437-963-128149	Sequence 128149,
45	43	47.3	1165	11	US-09-837-961-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-10-092-750-48
; Sequence 48, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; TYPE: PRT
; LENGTH: 18
; ORGANISM: Homo sapiens
US-10-092-750-48

Query Match 100.0%; Score 91; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 7.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFSVLLRYLADNPLPGGS 18

Db 1 LFSVLLRYLADNPLPGGS 18

RESULT 2
US-10-428-961-22
; Sequence 22, Application US/10428961
; Publication No. US20030237111A1
; GENERAL INFORMATION:

; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rugar, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; TITLE OF INVENTION: Polynucleotides, Compositions, and Methods of Use (Amended)
; FILE REFERENCE: MEC0201-1
; CURRENT APPLICATION NUMBER: US/10/428,961
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-428-961-22

Query Match 52.7%; Score 48; DB 14; Length 1168;
Best Local Similarity 61.5%; Pred. No. 76;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
||:|:|:|:|:
Db 41 LLQFLNNFVPGG 53

RESULT 3
US-10-614-524-4
; Sequence 4, Application US/10614524
; Publication No. US20040016020A1
; GENERAL INFORMATION:
; APPLICANT: Arnaut, Greta
; APPLICANT: Boete, Annemie
; APPLICANT: Damme, Nicole
; APPLICANT: Mathieu, Eva
; APPLICANT: Vanneste, Stijn
; APPLICANT: Van Rie, Jeroen
; TITLE OF INVENTION: Insecticidal proteins from Bacillus thuringiensis.
; FILE REFERENCE: NEWTSUS2
; CURRENT APPLICATION NUMBER: US/10/614,524
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US/09/739,243
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/173387
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-614-524-4

Query Match 52.7%; Score 48; DB 15; Length 1168;
Best Local Similarity 61.5%; Pred. No. 76;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
||:|:|:|:|:
Db 41 LLQFLNNFVPGG 53

RESULT 4
US-10-425-115-281976
; Sequence 281976, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 281976
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_20262C.1.pep
US-10-425-115-281976

Query Match 50.5%; Score 46; DB 17; Length 97;
Best Local Similarity 81.8%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 YLADNPLPGGS 18
||:|:|:|:|:
Db 42 YLFDIPLPGGS 52

RESULT 5
US-10-437-963-173457
; Sequence 173457, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 173457
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71493C.1.pep
US-10-437-963-173457

Query Match 50.5%; Score 46; DB 16; Length 318;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNPLPGGS 18
||:|:|:|:|:
Db 86 LLDLHLRHADKVAPOGA 103

RESULT 6
US-10-425-115-259841
; Sequence 259841, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants

```
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 259841
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3632-026-F12_FLI.pep
US-10-425-115-259841

Query Match          50.5%; Score 46; DB 17; Length 619;
Best Local Similarity 52.9%; Pred. NO. 82;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNLFPGG 17
Db 100 LLSGILLYLGSTYLPFG 116

RESULT 7
US-10-425-114-37487
; Sequence 37487, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37487
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3593-007-E1_FLI.pep
US-10-425-114-37487

Query Match          50.5%; Score 46; DB 15; Length 622;
Best Local Similarity 52.9%; Pred. NO. 82;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNLFPGG 17
Db 103 LLSGILLYLGSTYLPFG 119

RESULT 8
US-10-425-114-60096
; Sequence 60096, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60096

; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 259841
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3632-026-F12_FLI.pep
US-10-425-114-60096

Query Match          50.5%; Score 46; DB 15; Length 622;
Best Local Similarity 52.9%; Pred. NO. 82;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNLFPGG 17
Db 103 LLSGILLYLGSTYLPFG 119

RESULT 9
US-10-425-114-72803
; Sequence 72803, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 72803
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4702-003-B5_FLI.pep
US-10-425-114-72803

Query Match          50.5%; Score 46; DB 15; Length 622;
Best Local Similarity 52.9%; Pred. NO. 82;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNLFPGG 17
Db 103 LLSGILLYLGSTYLPFG 119

RESULT 10
US-10-424-599-147749
; Sequence 147749, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 147749
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(71)
; OTHER INFORMATION: unsure at all Xaa locations
```

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MET3847_104438C.1.pep
US-10-424-599-147749

Query Match 49.5%; Score 45; DB 15; Length 71;
Best Local Similarity 69.2%; Pred. NO. 12;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 LRYLADNFPVGG 18
Db 16 LSHLADHFLPWGS 28

RESULT 11
US-09-918-485-2
; Sequence 2, Application US/09918485
; Publication No. US20030115628A1
; GENERAL INFORMATION:
; APPLICANT: SANCHIS, Vincent
; LERECIUS, Didier
; MENOUE, Ghislaine
; LECADDET, Marguerite-Marie
; MARTOURET, Daniel
; DEMONDER, Raymond
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR
; POLYPEPTIDES ENDOWED WITH A LARVICIDAL ACTIVITY TOWARDS
; LEPIDOPTERA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/918,485
; FILING DATE: 25-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,551
; FILING DATE: 05-JUN-1995
; APPLICATION NUMBER: US 08/251,652
; FILING DATE: 31-MAY-1994
; APPLICATION NUMBER: US 07/458,754
; FILING DATE: 11-DEC-1989
; APPLICATION NUMBER: EP 88 401 121.4
; FILING DATE: 06-MAY-1988
; APPLICATION NUMBER: FR 87 08090
; FILING DATE: 10-JUN-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: HUNTINGTON, R. D.
; REGISTRATION NUMBER: 27,903
; REFERENCE/DOCKET NUMBER: 010830-073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 823 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-918-485-2

Query Match 49.5%; Score 45; DB 10; Length 823;
Best Local Similarity 53.8%; Pred. NO. 1.6e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LRYLADNFPVGG 17
Db 41 LVQFLVSNFVPGG 53

RESULT 12
US-09-826-660-21
; Sequence 21, Application US/09826660
; Patent No. US20010028940A1
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: WA-714XC2D1
; CURRENT APPLICATION NUMBER: US/09/826,660
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Toxin encoded by synthetic B.t. gene
US-09-826-660-21

Query Match 49.5%; Score 45; DB 9; Length 1163;
Best Local Similarity 53.8%; Pred. NO. 2.3e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LRYLADNFPVGG 17
Db 41 LVQFLVSNFVPGG 53

RESULT 13
US-09-972-175-2
; Sequence 2, Application US/09972175
; Publication No. US20030101482A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; Gilmer, Amy Jelen
; Mettus, Anne-Marie Light
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING
; LEPIDOPTERAN-ACTIVE DELTA-ENDOTOXINS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/972,175
; FILING DATE: 05-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/337,635
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:206
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1189 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-972-175-2

Query Match 49.5%; Score 45; DB 10; Length 1189;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LLRYLADNLFPGG 17
|::| |::|
Db 41 LVQFLVSNFVPGG 53

RESULT 14

US-09-972-175-4
Sequence 4, Application US/09972175
Publication No. US20030101482A1
GENERAL INFORMATION:
APPLICANT: Baum, James A.
Gilmer, Amy Jelen
Metzger, Anne-Marie Light
TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING
LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS

NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/972,175
FILING DATE: 05-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/337,635
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:206
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1189 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-972-175-4

Query Match 49.5%; Score 45; DB 10; Length 1189;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LLRYLADNLFPGG 17
|::| |::|
Db 41 LVQFLVSNFVPGG 53

RESULT 15

US-09-972-175-6
Sequence 6, Application US/09972175
Publication No. US20030101482A1
GENERAL INFORMATION:
APPLICANT: Baum, James A.
Gilmer, Amy Jelen
Metzger, Anne-Marie Light
TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING
LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS

NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/972,175
FILING DATE: 05-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/337,635
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: MECO:206
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1189 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-972-175-6

Query Match 49.5%; Score 45; DB 10; Length 1189;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LLRYLADNLFPGG 17
|::| |::|
Db 41 LVQFLVSNFVPGG 53

Search completed: November 11, 2004, 02:43:04
Job time : 21.8689 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 4.41509 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-48

Perfect score: 91

Sequence: 1 LFSVLLRYLADNPLPGGS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	50.5	664	2 AB1136	NADH flavin oxidor
2	46	50.5	664	2 AD1494	NADH flavin oxidor
3	45	49.5	570	2 A75201	hypothetical prote
4	45	49.5	570	2 E71234	hypothetical prote
5	45	49.5	655	2 JC7140	protoxin - Bacillu
6	45	49.5	678	2 B95968	probable NADH-depe
7	45	49.5	823	2 S04181	parasporeal crystal
8	45	49.5	1106	2 T31742	hypothetical prote
9	45	49.5	1189	2 S00944	parasporeal crystal
10	44	48.4	473	2 F70513	probable PPE prote
11	43	47.3	268	2 G71472	hypothetical prote
12	43	47.3	1160	2 S32647	parasporeal crystal
13	43	47.3	1165	2 S11446	parasporeal crystal
14	42.5	46.7	2113	2 G31286	probable RNA helic
15	42.5	46.7	2113	2 C86128	probable helicase
16	42	46.2	238	2 A83224	probable ATP-bindi
17	42	46.2	947	2 T00340	hypothetical prote
18	42	46.2	1345	2 S55669	tegument protein 7
19	42	46.2	1679	2 T50091	yeast Ecm29 cell w
20	41	45.1	149	2 F65169	O149 protein - Esc
21	41	45.1	149	2 D86052	hypothetical prote
22	41	45.1	149	2 B91206	hypothetical prote
23	41	45.1	181	2 AG1853	sucrose-phosphatas
24	41	45.1	319	2 B70194	conserved hypotet
25	41	45.1	356	2 T51105	glucose-1-phosphat
26	41	45.1	385	2 T05049	hypothetical prote
27	41	45.1	612	2 G64678	NADH2 dehydrogenas
28	41	45.1	612	2 E71839	NADH2 dehydrogenas
29	41	45.1	622	2 G81981	probable lipopolys

ALIGNMENTS

RESULT 1

AB1136

NADH flavin oxidoreductase homolog lmo0489 [imported] - Listeria monocytogenes (strain
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AB1136
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, I.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; P
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlanc
A;Title: Comparative Genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1136
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-664 <GLA>
A;Cross-references: UNIPROT:Q8Y9N6; GB:NC_003210; PIDN:CAC98568.1; PID:gl6409865; GSPDE
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo0489
C;Superfamily: Methylophilus methylotrophus WJAL trimethylamine dehydrogenase

Query Match 50.5%; Score 46; DB 2; Length 664;

Best Local Similarity 47.1%; Pred. No. 13;

Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY	2	FSVLLRYLADNPLPGGS	18
DB	218	FPVILRFSADEFTEGGN	234

RESULT 2

AD1494

NADH flavin oxidoreductase homolog lin0492 [imported] - Listeria innocua (strain Clip11
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AD1494
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, I.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; P
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlanc
A;Title: Comparative Genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AD1494
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-664 <GLA>
A;Cross-references: UNIPROT:Q82EG8; GB:AL592022; PIDN:CAC95724.1; PID:gl6412932; GSPDB

A;Experimental source: strain Clp11262

C;Genetics:

A;Gene: lin0492

C;Superfamily: Methylophilus methylotrophus W3A1 trimethylamine dehydrogenase

Query Match 50.5%; Score 46; DB 2; Length 664;

Best Local Similarity 47.1%; Pred. No. 13; Indels 0; Gaps 0;

Matches 8; Conservative 4; Mismatches 5;

QY 2 FSVLLRYLADNPLPGGS 18

|||:|||||:

Db 218 FSVLRFSADETEGNG 234

RESULT 3

A75201

hypothetical protein PAB0082 - Pyrococcus abyssi (strain Orsay)

C;Species: Pyrococcus abyssi

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: A75201

R;anonymous, Genoscope

A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A;Reference number: A75001

A;Accession: A75201

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-570 <RAW>

A;Cross-references: UNIPROT:Q9V2E2; GB:AJ248283; GB:AL096836; NID:gs457433; PIDN:CA54905

A;Experimental source: strain Orsay

C;Genetics:

A;Gene: PAB0082

Query Match 49.5%; Score 45; DB 2; Length 570;

Best Local Similarity 80.0%; Pred. No. 16; Indels 1; Gaps 0;

Matches 8; Conservative 1; Mismatches 1;

QY 6 LRYLADNPLP 15

|||||:

Db 211 LRYLVENFLP 220

RESULT 4

E71234

hypothetical protein PH0135 - Pyrococcus horikoshii

C;Species: Pyrococcus horikoshii

C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004

C;Accession: E71234

R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Accession: E71234

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-570 <RAW>

A;Cross-references: UNIPROT:Q57875; GB:AP000001; NID:g3236128; PIDN:BAA29204.1; PID:g325

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBank

C;Genetics:

A;Gene: PH0135

Query Match 49.5%; Score 45; DB 2; Length 570;

Best Local Similarity 80.0%; Pred. No. 16; Indels 1; Gaps 0;

Matches 8; Conservative 1; Mismatches 1;

QY 6 LRYLADNPLP 15

|||||:

Db 211 LRYLVENFLP 220

RESULT 5

JC7140

prototoxin - Bacillus thuringiensis

N;Alternate names: cryIC protein

C;Species: Bacillus thuringiensis

C;Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 11-May-2000

C;Accession: JC7140

R;Christov, N.K.; Imaishi, H.; Ohkawa, H.

Biosci. Biotechnol. Biochem. 63, 1433-1444, 1999

A;Title: Green-tissue-specific expression of a reconstructed cryIC gene encoding the ac

ptera litura.

A;Reference number: JC7140; MUID:99430790; PMID:10501003

A;Accession: JC7140

A;Molecule type: DNA

A;Residues: 1-655 <CHR>

A;Cross-references: EMBL:X96682

C;Genetics:

A;Gene: cryIC

C;Superfamily: parasporal crystal protein

Query Match 49.5%; Score 45; DB 2; Length 655;

Best Local Similarity 53.8%; Pred. No. 18; Indels 0; Gaps 0;

Matches 7; Conservative 4; Mismatches 2;

QY 5 LRYLADNPLPGG 17

|||||:

Db 41 LVQFLVSNFVPGG 53

RESULT 6

B95968

probable NADH-dependent oxidase protein stcD2 [imported] - Sinorhizobium meliloti (stra

C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C;Accession: B95968

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing end

A;Reference number: A95842; MUID:21396508; PMID:11481431

A;Accession: B95968

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-678 <KUR>

A;Cross-references: UNIPROT:O87278; GB:AL591985; PIDN:CA49410.1; PID:g15140896; GSPDB:

A;Experimental source: strain 1021, megaplasmid pSymb

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure

heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: stcD2; SMD21570

A;Genome: plasmid

C;Superfamily: Methylophilus methylotrophus W3A1 trimethylamine dehydrogenase

Query Match 49.5%; Score 45; DB 2; Length 678;

Best Local Similarity 52.9%; Pred. No. 19; Indels 0; Gaps 0;

Matches 9; Conservative 3; Mismatches 5;

QY 2 FSVLLRYLADNPLPGGS 18

|||:|||||:

Db 215 FILGVRYTADSCUPGGT 231

RESULT 7

S04181

Parasporal crystal protein - Bacillus thuringiensis (strain aizawai 7.29) (fragment)

N;Alternate names: delta-endotoxin

C;Species: Bacillus thuringiensis

C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004

C:Accession: S04181
R:Sanchez, V.; Lereclus, D.; Menou, G.; Chauvaux, J.; Guo, S.; Lecadet, M.M.
Mol. Microbiol. 3, 229-238, 1989
A:Title: Nucleotide sequence and analysis of the N-terminal coding region of the Spodoptera frugiperda delta-endotoxin
A:Reference number: S04181; MUID:89343627; PMID:2548060
A:Accession: S04181
A:Molecule type: DNA
A:Residues: 1-823 <SAS>
A:Cross-references: UNIPROT:P05518; EMBL:X13620; NID:g40355; PIDN:CAA31951.1; PID:g40356
C:Genetics:
A:Gene: bta
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 49.5%; Score 45; DB 2; Length 823;
Best Local Similarity 53.8%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LRLYLADNLFPGG 17
|::| |::|
Db 41 LVQFLVSNFVPGG 53

RESULT 8
T31742
Hypothetical protein C05C8.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31742
R:Sammons, L.; Wohldmann, P.
Submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid C05C8.
A:Reference number: Z21078
A:Accession: T31742
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1106 <SAM>
A:Cross-references: UNIPROT:O16310; EMBL:AF016430; PIDN:AAB65371.1; GSPDB:GN00023; CESP:
A:Experimental source: strain Bristol N2; clone C05C8
C:Genetics:
A:Gene: CESP:C05C8.4
A:Map position: 5
A:Introns: 25/3; 78/3; 117/1; 245/1; 591/1; 787/1; 1008/2

Query Match 49.5%; Score 45; DB 2; Length 1106;
Best Local Similarity 55.6%; Pred. No. 32;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNLFPGGS 18
|::| |::|
Db 738 LFEALSKVQADNFSGGS 755

RESULT 9
S00944
Parasporal crystal protein cryIcal - Bacillus thuringiensis (strain entomocidus 60.5)
C:Species: Bacillus thuringiensis
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S00944
R:Honee, G.; van der Salm, T.; Visser, B.
Nucleic Acids Res. 16, 6240, 1988
A:Title: Nucleotide sequence of crystal protein gene isolated from B. thuringiensis subsp. israelensis
A:Reference number: S00944; MUID:88289380; PMID:3399402
A:Accession: S00944
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-1189 <HON>
A:Cross-references: UNIPROT:P05518; EMBL:X07518; NID:g40293; PIDN:CAA30396.1; PID:g40294
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 49.5%; Score 45; DB 2; Length 1189;
Best Local Similarity 53.8%; Pred. No. 34;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 LRLYLADNLFPGG 17
|::| |::|
Db 41 LVQFLVSNFVPGG 53

RESULT 10
F70513
Probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: F70513
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70513
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-473 <COL>
A:Cross-references: GB:297559; GB:AL123456; NID:g3261820; PIDN:CAB10718.1; PID:g328403
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: PPE

Query Match 48.4%; Score 44; DB 2; Length 473;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 FSVLLRYLADNLFPG 16
|::| |::|
Db 204 FDELRYLAVELPG 218

RESULT 11
G71472
Hypothetical protein CT764 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: G71472
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, S.; Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, S.
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: G71472
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-268 <ARN>
A:Cross-references: UNIPROT:O84769; GB:AE001349; GB:AE001273; NID:g3329226; PIDN:AAC68:
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: CT764
C:Superfamily: Chlamydia trachomatis hypothetical protein CT764

Query Match 47.3%; Score 43; DB 2; Length 268;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 LRLYLADNLFPG 16
|::| |::|
Db 223 LFRFRXDSFLPG 234

RESULT 12
S32647
Parasporal crystal protein cryIb1 - Bacillus thuringiensis
C:Species: Bacillus thuringiensis
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S32647

R.Lambert, B.
 submitted to the EMBL Data Library, April 1993
 A:Reference number: S32645
 A:Accession: S32647
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-1160 <LAM>
 A:Cross-references: UNIPROT:Q45747; EMBL:222511; NID:G295863; PIDN:CAA80234.1; PID:G29586
 C:Superfamily: parasporal crystal protein
 C:Keywords: delta-endotoxin

Query Match 47.3%; Score 43; DB 2; Length 1160;
 Best Local Similarity 53.8%; Pred. No. 73;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
 : : |||
 Db 41 LINFLYSNFVPGG 53

RESULT 13

S11446
 parasporal crystal protein cryIba1 - Bacillus thuringiensis
 N:Alternate names: parasporal crystal protein cryID
 C:Species: Bacillus thuringiensis
 C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C:Accession: S11446

R.Hoeite, H.; Soetaert, P.; Janssens, S.; Peferoen, M.

Nucleic Acids Res. 18, 5545, 1990
 A:Title: Nucleotide sequence and deduced amino acid sequence of a new Lepidoptera-specific
 A:Reference number: S11446; MUID:91016842; PMID:2215728
 A:Accession: S11446
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1165 <HOB>
 A:Cross-references: UNIPROT:P19415; EMBL:X54160; NID:G40279; PIDN:CAA38099.1; PID:G40280
 C:Superfamily: parasporal crystal protein
 C:Keywords: delta-endotoxin

Query Match 47.3%; Score 43; DB 2; Length 1165;
 Best Local Similarity 53.8%; Pred. No. 73;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 LLRYLADNPLPGG 17
 : : |||
 Db 41 LINFLYSNFVPGG 53

RESULT 14

G91286
 probable RNA helicase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C:Accession: G91286
 R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: G91286
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2113 <HAY>
 A:Cross-references: UNIPROT:Q8XC71; GB:BA000007; PIDN:BA838686.1; PID:G13364741; GSPDB:C
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECs5263

Query Match 46.7%; Score 42.5; DB 2; Length 2113;
 Best Local Similarity 62.5%; Pred. No. 1.7e+02;
 Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 2 FSVLLRYLADNPLPGG 17

Db 1850 FALLLDYLAD---PGG 1862
 : : ||| ||| |||

RESULT 15

C86128
 Probable helicase Z5901 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C:Date: 18-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: C86128
 R.Ferna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: C86128
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2113 <STO>
 A:Cross-references: UNIPROT:Q8XC71; GB:AE005174; NID:G12519309; PIDN:AAG59487.1; GSPDB:C
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: Z5901

Query Match 46.7%; Score 42.5; DB 2; Length 2113;
 Best Local Similarity 62.5%; Pred. No. 1.7e+02;
 Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 2 FSVLLRYLADNPLPGG 17
 : : ||| ||| |||
 Db 1850 FALLLDYLAD---PGG 1862

Search completed: November 10, 2004, 14:52:17
 Job time : 5.41509 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 23.9774 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-48

Perfect score: 91

Sequence: 1 LFSVLLRYLADNPLPGGS 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	52.7	782	2 Q6PYW6	Q6PYW6 bacillus th
2	48	52.7	782	2 AAS93799	AAS93799 bacillus
3	48	52.7	1167	1 C1JA BACTU	Q45738 bacillus th
4	48	52.7	1170	1 C1JB BACTU	Q45716 bacillus th
5	47	51.6	604	2 Q3678	Q23678 arabidopsis
6	47	51.6	622	2 Q39045	Q39045 arabidopsis
7	47	51.6	623	2 Q39047	Q39047 arabidopsis
8	47	51.6	1169	1 C1GB BACTZ	Q92426 bacillus th
9	47	51.6	1169	2 Q8GHE8	Q8GHE8 bacillus th
10	46	50.5	280	2 Q9KIB5	Q9KIB5 azotobacter
11	46	50.5	282	2 Q89P21	Q89P21 bradyrhizob
12	46	50.5	354	2 Q7MT25	Q7MT25 streptomyce
13	46	50.5	606	2 Q7MBG3	Q7MBG3 wolnelia
14	46	50.5	664	2 Q32E88	Q32E88 listeria in
15	46	50.5	664	2 Q8Y9N6	Q8Y9N6 listeria in
16	46	50.5	664	2 Q723G0	Q723G0 listeria mo
17	46	50.5	664	2 AAT03301	AAT03301 listeria
18	46	50.5	1246	2 Q77CC8	Q77CC8 bovine herp
19	46	50.5	1246	2 Q89604	Q89604 bovine herp
20	46	50.5	1250	2 Q8X242	Q8X242 bovine herp
21	46	50.5	1250	2 AAR86131	AAR86131 bovine he
22	45	49.5	269	2 Q6FZX7	Q6FZX7 bartonella
23	45	49.5	354	2 Q6BS12	Q6BS12 debaryomyce
24	45	49.5	485	2 Q8H633	Q8H633 oryza sativ
25	45	49.5	570	2 Q57875	Q57875 pyrococcus
26	45	49.5	570	2 Q9V2E2	Q9V2E2 pyrococcus
27	45	49.5	570	2 Q8U4L1	Q8U4L1 pyrococcus
28	45	49.5	678	1 STCD_RHIME	Q87278 rhizobium m
29	45	49.5	764	2 Q6FEW8	Q6FEW8 acinetobact
30	45	49.5	803	2 Q8IDF9	Q8IDF9 plasmodium
31	45	49.5	1106	2 O16310	O16310 caenorhabdi

32	45	49.5	1189	1 C1CA BACTE	P05518 bacillus th
33	45	49.5	1189	2 Q6YNB8	Q6YNB8 bacillus th
34	45	49.5	1189	2 Q9L877	Q9L877 bacillus th
35	45	49.5	1189	2 AAM00264	AAM00264 bacillus
36	45	49.5	1189	2 AAL79362	AAL79362 bacillus
37	44	48.4	143	2 Q98WQ3	Q98WQ3 rhizobium 1
38	44	48.4	310	2 Q6DFM6	Q6DFM6 xenopus tro
39	44	48.4	454	2 Q7VEQ6	Q7VEQ6 mycobacteri
40	44	48.4	473	2 Q79FH3	Q79FH3 mycobacteri
41	44	48.4	473	2 CAES5457	CAES5457 mycobacte
42	44	48.4	580	2 Q7Q063	Q7Q063 anopheles g
43	44	48.4	751	2 Q7NDW2	Q7NDW2 gloeobacter
44	43.5	47.8	385	1 CYB_EPTBU	Q942d1 eptatretus
45	43	47.3	181	2 Q8GKT8	Q8GKT8 bacillus ce

ALIGNMENTS

RESULT 1

Q6PYW6 PRELIMINARY; PRT; 782 AA.

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cyl type crystal protein (fragment).
OS Bacillus thuringiensis (subsp. kenya).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=33930;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K3;
RA Li M.S., Choi J.Y., Roh J.Y., Shim H.J., Boo K.S., Je Y.H.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY570736; AAS93799.1; -
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal bind like.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
FT NON_TER 782
SQ SEQUENCE 782 AA; 89009 MW; 36990BD477860DB9 CRC64;

Query Match 52.7%; Score 48; DB 2; Length 782;
Best Local Similarity 61.5%; Pred. No. 39;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 5 LLEYLADNPLPGG 17
||:|:|:|:|:|:|
Db 35 LLQFLNNFVPGG 47

RESULT 2

AAS93799 PRELIMINARY; PRT; 782 AA.

DT 26-APR-2004 (TrEMBLrel. 27, Created)
DT 26-APR-2004 (TrEMBLrel. 27, Last sequence update)
DT 26-APR-2004 (TrEMBLrel. 27, Last annotation update)
DE Cyl type crystal protein (fragment).
OS Bacillus thuringiensis (subsp. kenya).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group; Bacillus thuringiensis.
OX NCBI_TaxID=33930;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K3;
RA Li M.S., Choi J.Y., Roh J.Y., Shim H.J., Boo K.S., Je Y.H.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY570736; AAS93799.1; -

```

AC Q45716;
AC 30-MAY-2000 (Rel. 39, Created)
AC 30-MAY-2000 (Rel. 39, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE PEPTIDICAL CRYSTAL PROTEIN cryIJB (Insecticidal delta-endotoxin
DE cryIJ(b)) (Crystalline entomocidial protoxin) (134 kDa crystal protein).
GN Name=cryIJB; Synonyms=cryIJ(b), cryETI;
GN Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxID=1428;
RN [1]
SEQUENCE FROM N.A.
STRAIN=NRLR B-18973 / EG5092;
RA von Tersch M.A., Gonzalez J.M. Jr.;
RA "Bacillus thuringiensis cryETI toxin gene and protein toxic to
RT lepidopteran insects";
RRL Parent number US5356623, 18-OCT-1994.
CC -! FUNCTION: Promotes colloid osmotic lysis by binding to the midgut
CC epithelial cells of many lepidopteran larvae. Toxic to Plutella
CC xylosteella.
CC
CC -! DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC
CC -! MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC
CC -! SIMILARITY: Belongs to the delta endotoxin family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
-----
DR EMBL; U31527; AAA98959.1; --
DR HSSP; P02965; ICIY.
DR InterPro; IPRO01178; Endotoxin.
DR InterPro; IPRO05638; endotoxin_C.
DR InterPro; IPRO05639; endotoxin_N.
DR InterPro; IPRO08979; Gal_bind_like.
DR Pfam; PF01944; Endotoxin_C_1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
DR Sporulation; Toxin.
KW SEQUENCE 1170 AA; 133553 MW; B053B9619B78DC19 CRC64;
SQ
Query Match 52.7%; Score 48; DB 1; Length 1170;
Best Local Similarity 61.5%; Pred. No. 59;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 5 LLRYLDNFLPGG 17
DB 41 LLQLLNFFVGG 53
||::|:|||
||::|:|||
RESULT 5
O23678 PRELIMINARY; PRT; 604 AA.
AC O23678;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CER1-like protein.
DE Name=T7123.9;
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
OX [1]
SEQUENCE FROM N.A.
RA Federspiel N.A., Conway A.B., Conway A.R., Davis K., Brendel V.,

```



```
RA Palm C.J., Au M., Araujo R., Chung E., Kurtz D.B., Buehler E.,
RA Dewar K., Peng J., Kim C., Li Y., Shinn P., Sun H., Oji O.,
RA Osborne B., Shen Y.K., Toriumi M., Vyotekala V., Yu G., Theologis A.,
RA Ecker J., Davis R.W.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; U89959; AAC24373.1; -; catalytic activity; IEA.
DR GO; GO:0003824; P:metabolic; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006087; Sterol_desat.
DR InterPro; IPR006088; Sterol_desatur.
DR Pfam; PF01598; Sterol_desat; 1.
SQ SEQUENCE 604 AA; 65334 MW; B0C46279AA473782 CRC64;

Query Match 51.6%; Score 47; DB 2; Length 604;
Best Local Similarity 55.6%; Pred. No. 4;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNPLPGGS 18
Db 100 IFNTLLMYLANIKLPGAS 117

RESULT 6
Q39046 PRELIMINARY; PRT; 622 AA.
ID Q39046
AC Q39046;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE CER1-like protein.
GN Name=CER1-like;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Deng M.D., Peng S., Lemieux B.;
RT "Genomic (Accession No. X95964) and cDNA (Accession No. X95965)
RT sequences of the CER1-like gene of Arabidopsis thaliana derived from a
RT Plant DNA/T-DNA insertion junction. (FGR96-019).";
RL Plant Physiol. 110:1436-1436(1996).
DR EMBL; X95964; CAA65199.1; -.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR InterPro; IPR006087; Sterol_desat.
DR InterPro; IPR006088; Sterol_desatur.
DR Pfam; PF01598; Sterol_desat; 1.
SQ SEQUENCE 622 AA; 71457 MW; DDEF0FALC6967493 CRC64;

Query Match 51.6%; Score 47; DB 2; Length 622;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNPLPGGS 18
Db 100 IFNTLLMYLANIKLPGAS 117

RESULT 7
Q39047 PRELIMINARY; PRT; 623 AA.
ID Q39047
AC Q39047;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE CER1-like protein.
GN Name=CER1-like;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Deng M.D., Peng S., Lemieux B.;
RT "Genomic (Accession No. X95964) and cDNA (Accession No. X95965)
RT sequences of the CER1-like gene of Arabidopsis thaliana derived from a
RT Plant DNA/T-DNA insertion junction. (FGR96-019).";
RL Plant Physiol. 110:1436-1436(1996).
DR EMBL; X95964; CAA65199.1; -.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR InterPro; IPR006087; Sterol_desat.
DR InterPro; IPR006088; Sterol_desatur.
DR Pfam; PF01598; Sterol_desat; 1.
SQ SEQUENCE 622 AA; 71457 MW; DDEF0FALC6967493 CRC64;

Query Match 51.6%; Score 47; DB 2; Length 622;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNPLPGGS 18
Db 100 IFNTLLMYLANIKLPGAS 117

RESULT 8
Q39048 PRELIMINARY; PRT; 1169 AA.
ID Q39048
AC Q39048;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pesticidal crystal protein cryIgb (Insecticidal delta-endotoxin
DE CryIgb) (crystalline entomocidal protoxin) (133 kDa crystal protein).
GN Name=cryIgb; Synonyms=cryIgb(b), cryH2;
OS Bacillus thuringiensis (subsp. wuhanensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxID=52024;
RN [1]
RP SEQUENCE FROM N.A.
RA Kuo W.-S., Lin J.-H., Tzeng C.-C., Kao S.-S., Chak K.-F.;
RT "Cloning of two new cry genes from Bacillus thuringiensis subsp.
RT wuhanensis strain.";
RL Curr. Microbiol. 40:227-232(2000).
DR EMBL; U70725; AAD10291.1; -.
DR HSP; P02965; ICIV.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
DR Spore; P03945; Endotoxin_N; 1.
KW Sporulation; Toxin.
```

```
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Deng M.D., Peng S., Lemieux B.;
RT "Genomic (Accession No. X95964) and cDNA (Accession No. X95965)
RT sequences of the CER1-like gene of Arabidopsis thaliana derived from a
RT Plant DNA/T-DNA insertion junction. (FGR96-019).";
RL Plant Physiol. 110:1436-1436(1996).
DR EMBL; X95965; CAA65200.1; -.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR InterPro; IPR006087; Sterol_desat.
DR InterPro; IPR006088; Sterol_desatur.
DR Pfam; PF01598; Sterol_desat; 1.
SQ SEQUENCE 623 AA; 71571 MW; BD896FP94EC5718C CRC64;

Query Match 51.6%; Score 47; DB 2; Length 623;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LFSVLLRYLADNPLPGGS 18
Db 100 IFNTLLMYLANIKLPGAS 117

RESULT 8
Q39049 PRELIMINARY; PRT; 1169 AA.
ID Q39049
AC Q39049;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pesticidal crystal protein cryIgb (Insecticidal delta-endotoxin
DE CryIgb) (crystalline entomocidal protoxin) (133 kDa crystal protein).
GN Name=cryIgb; Synonyms=cryIgb(b), cryH2;
OS Bacillus thuringiensis (subsp. wuhanensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxID=52024;
RN [1]
RP SEQUENCE FROM N.A.
RA Kuo W.-S., Lin J.-H., Tzeng C.-C., Kao S.-S., Chak K.-F.;
RT "Cloning of two new cry genes from Bacillus thuringiensis subsp.
RT wuhanensis strain.";
RL Curr. Microbiol. 40:227-232(2000).
DR EMBL; U70725; AAD10291.1; -.
DR HSP; P02965; ICIV.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
DR Spore; P03945; Endotoxin_N; 1.
KW Sporulation; Toxin.
```

```
SQ SEQUENCE 1169 AA; 132904 MW; D1EPC1508A9B10BD CRC64;
Query Match 51.6%; Score 47; DB 1; Length 1169;
Best Local Similarity 57.1%; Pred. No. 87;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 LRLYLADNPLPGGS 18
: | : | : | : | : | : |
Db 40 ITRLILENPLPGGS 53

RESULT 9
Q8GHE8 PRELIMINARY; PRT; 1169 AA.
AC Q8GHE8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Delta-endotoxin Cry1.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=B-Pr-88;
RA Li C., Zhang J., Li G., Huang D.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF288683; AAO13756.1; -
DR HSSP; P02965; ICIV.
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
SQ SEQUENCE 1169 AA; 132990 MW; 07C9D6D180F3DB6D CRC64;

Query Match 51.6%; Score 47; DB 2; Length 1169;
Best Local Similarity 57.1%; Pred. No. 87;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 LRLYLADNPLPGGS 18
: | : | : | : | : | : |
Db 40 ITRLILENPLPGGS 53

RESULT 10
Q9KIB5 PRELIMINARY; PRT; 280 AA.
AC Q9KIB5
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Ampe.
GN Name=ampE;
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
[1]
RN SEQUENCE FROM N.A.
RC MEDLINE=20398165; PubMed=10940024;
RA Nunez C., Moreno S., Cardenas L., Soberon-Chavez G., Espin G.;
RX "Inactivation of the ampE operon increases transcription of algD and
RT affects morphology and encystment of Azotobacter vinelandii.";
RL J. Bacteriol. 182:4829-4835(2000).
DR EMBL; AF237388; AAF73951.1; -
DR SEQUENCE 280 AA; 31498 MW; 230869401568DDA6 CRC64;

SQ
```

```
Query Match 50.5%; Score 46; DB 2; Length 280;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FSVLLRYLADNPLP 15
: | : | : | : | : | : |
Db 211 VPRVLLRYLSREVP 225

RESULT 11
Q89PZ1 PRELIMINARY; PRT; 282 AA.
AC Q89PZ1
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ABC transporter permease protein.
GN OrderedLocustNames=blf3339;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiuni T.,
RX Sasanoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
CC -1- FUNCTION: Part of a binding-protein-dependent transport system.
CC -1- Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL; AP005947; BAC48604.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp_1; 1.
DR PROSITE; PS00928; ABC_TM1; 1.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 282 AA; 30199 MW; 1AFE04BB6512BDF4 CRC64;

Query Match 50.5%; Score 46; DB 2; Length 282;
Best Local Similarity 64.3%; Pred. No. 30;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 FSVLLRYLADNPLP 15
: | : | : | : | : | : |
Db 185 FSVLLRYLADNPLP 198

RESULT 12
Q7WT25 PRELIMINARY; PRT; 354 AA.
AC Q7WT25
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NRP-1-glucose synthase.
GN Name=NRP-1;
OS Streptomyces sp. AM-7161.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=221710;
[1]
RN SEQUENCE FROM N.A.
```

RC STRAIN=AM-7161;
RX MEDLINE=22739958; PubMed=12855716;
RA Ichinose K., Ozawa M., Ito K., Kunieda K., Ebizuka Y.;
RT "Cloning, sequencing and heterologous expression of the medermycin
RT biosynthetic gene cluster of Streptomyces sp. AM-7161: towards
RT comparative analysis of the benzochromanone gene clusters.";
RL Microbiology 149:1633-1645(2003).
CC -!- CATALYTIC ACTIVITY: dTTP + alpha-D-glucose 1-phosphate =
CC diphosphate + dTDP-glucose.
CC -!- SIMILARITY: Belongs to the glucose-1-phosphate
CC thymidyllyltransferase family.
DR EMBL: AB103463; BAC79029.1; -
DR GO: GO:0008879; F:glucose-1-phosphate thymidyltransferase a. . .; IEA.
DR GO: GO:0017000; P:antibiotic biosynthesis; IEA.
DR InterPro: IPR005908; GTP thy trans s.
DR InterPro: IPR005835; NTP transferase.
DR Pfam: PF00483; NTP transferase; 1.
DR TIGRFAMs: TIGR01208; xmla long; 1.
KW Kinase; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 354 AA; 37083 MW; 80D2BCF931ABAB51 CRC64;

Query Match 50.5%; Score 46; DB 2; Length 354;
Best Local Similarity 80.0%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 VLADNPLPGG 17
DB 104 YLGDNPLPGG 113

RESULT 13
Q7MBG3 PRELIMINARY; PRT; 606 AA.
AC Q7MBG3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE PUTATIVE ASPARAGINE SYNTHETASE (GLUTAMINE-HYDROLYZ (EC 6.3.5.4)).
GN Name=PA11605; OrderedLocusNames=NS2192;
OS Wolinella succinogenes.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Wolinella.
OX NCBI_TaxID=844;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSMZ 1740;
RX MEDLINE=22882897; PubMed=14500908;
RA Baat C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimmek O.,
RA Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,
RA Meyer F., Lederer H., Schuster S.C.;
RT "Complete genome sequence and analysis of Wolinella succinogenes";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).
DR EMBL: BX571662; CAE11183.1; -
DR GO: GO:0004065; P:asparagine synthase (glutamine-hydrolyzing). . .; IEA.
DR GO: GO:0016874; P:ligase activity; IEA.
DR GO: GO:0006529; P:asparagine biosynthesis; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR001962; Asn synthase.
DR InterPro: IPR006426; Asn synth_AEB.
DR Pfam: PF00733; Asn_synthase; 1.
DR Pfam: PF00310; GARase_2; 1.
DR TIGRFAMs: TIGR01536; asn synth_AEB; 1.
DR PROSITE: PS00443; GATASE_TYPE_II; UNKNOWN_1.
KW Complete proteome; Ligase.
SQ SEQUENCE 606 AA; 71101 MW; 4C01CD1B7F4DDDD16 CRC64;

Query Match 50.5%; Score 46; DB 2; Length 606;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 VLLRYLADNPLPGGS 18
DB 104 YLGDNPLPGG 113

Db 184 VILPYLARGFLDGG 198

RESULT 14
Q92EG8 PRELIMINARY; PRT; 664 AA.
AC Q92EG8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lmo0492 protein.
GN OrderedLocusNames=lmo0492;
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Etian K.-D., Esibi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkut G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- COFACTOR: FAD (By similarity).
DR EMBL: AL596165; CAC95724.1; -
DR PIR: AD1494; AD1494.
DR HSP: P71278; LGVR.
DR ListList; LIN0492; -
DR GO: GO:0015036; F:disulfide oxidoreductase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR000759; Admrx_reductase.
DR InterPro: IPR001327; FAD pyr_redox.
DR InterPro: IPR001155; NAD_BS.
DR InterPro: IPR001155; Oxidored FMN.
DR InterPro: IPR00103; Pyridine_redox_2.
DR Pfam: PF00724; Oxidored FMN; 1.
DR Pfam: PF00070; Pyr_redox; 1.
DR PRINTS: PR00419; ADXRDTASE.
DR PRINTS: PR00368; FADPNR.
DR PRINTS: PR00469; PNDRTASEII.
DR ProDom: PD00139; FAD pyr_redox; 1.
KW Complete proteome; FAD; Flavoprotein; Oxidoreductase.
SQ SEQUENCE 664 AA; 72912 MW; 57CA1857181B205D CRC64;

Query Match 50.5%; Score 46; DB 2; Length 664;
Best Local Similarity 47.1%; Pred. No. 72;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 FSVLLRYLADNPLPGGS 18
DB 218 FPIVLFPSADETEGCG 234

RESULT 15
ID Q8Y9N6 PRELIMINARY; PRT; 664 AA.
AC Q8Y9N6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE lmo0489 protein.
GN OrderedLocusNames=lmo0489;
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OX NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EGD-e / Serovar 1/2a;
 RX MEDLINE=2157279; PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Rain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Klett J., Kuhn M., Kunst F., Kurapkat G.,
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 RT "Comparative genomics of *Listeria species*,"
 RJ Science 294:849-852(2001).
 CC -!- COFACTOR: FAD (By similarity).
 DR EMBL; AL591975; CAC98568.1; -.
 DR PIR; AB1136; AB1136.
 DR HSSP; P71278; 1GVR.
 DR ListList; LMO0489; -.
 DR GO; GO:0015036; Fdisulfide oxidoreductase activity; IEA.
 DR GO; GO:0008118; P/electron transport; IEA.
 DR InterPro; IPR000759; Adrndx_reductase.
 DR InterPro; IPR001327; FAD_PyT_redox.
 DR InterPro; IPR000205; NAD_BS.
 DR InterPro; IPR001155; Oxidored FMN.
 DR InterPro; IPR000103; Pyridine_redox_2.
 DR Pfam; PF00724; Oxidored FMN; 1.
 DR Pfam; PF00070; PyT_redox; 1.
 DR PRINTS; PR00419; ADXRDTASE.
 DR PRINTS; PR00368; FADPNR.
 DR PRINTS; PR00469; PNDRTASEII.
 DR ProDom; PD000139; FAD_PyT_redox; 1.
 KW Complete proteome; FAD; Flavoprotein; Oxidoreductase.
 SQ SEQUENCE 664 AA; 72854 MW; E45852E49C3EFCC CRC64;

Query Match 50.5%; Score 46; DB 2; Length 664;
 Best Local Similarity 47.1%; Pred. No. 72;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 FSVLLRYLADNPLPGGS 18
 DB 218 PFIVLRFSADEFTGGN 234

Search completed: November 10, 2004, 14:50:06
 Job time : 25.9774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 6.58868 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-49
Perfect score: 102
Sequence: 1 DWQLLCKLLWKIDNPGI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	46.5	45.6	4 US-09-621-976-6534	Sequence 6534, Ap
2	46.5	45.6	55 4 US-09-621-976-6991	Sequence 6991, Ap
3	46.5	45.6	189 4 US-09-543-681A-7788	Sequence 7788, Ap
4	45.5	44.6	280 2 US-08-464-517-6	Sequence 6, Appl
5	45.5	44.6	280 3 US-08-463-772-6	Sequence 6, Appl
6	45.5	44.6	291 5 PCT-US93-05000-6	Sequence 6, Appl
7	45.5	44.6	292 2 US-08-464-517-23	Sequence 23, Appl
8	45.5	44.6	292 2 US-08-246-361A-23	Sequence 23, Appl
9	45.5	44.6	292 2 US-08-246-361A-23	Sequence 23, Appl
10	45.5	44.6	292 3 PCT-US93-05000-23	Sequence 23, Appl
11	45.5	44.6	292 5 PCT-US93-05000-23	Sequence 23, Appl
12	45	44.1	20 4 US-09-556-877-246	Sequence 246, App
13	45	44.1	20 4 US-09-556-877-247	Sequence 247, App
14	45	44.1	20 4 US-09-620-412C-246	Sequence 246, App
15	45	44.1	20 4 US-09-598-419-246	Sequence 246, App
16	45	44.1	20 4 US-09-598-419-247	Sequence 247, App
17	45	44.1	20 4 US-09-598-419-247	Sequence 247, App
18	45	44.1	556 4 US-09-198-452A-596	Sequence 596, App
19	44	43.1	85 4 US-09-513-999C-7713	Sequence 7713, Ap
20	44	43.1	608 4 US-09-325-932A-179	Sequence 179, Ap
21	43.5	42.6	174 2 US-08-464-517-50	Sequence 50, Appl
22	43.5	42.6	174 2 US-08-246-361A-50	Sequence 50, Appl
23	43.5	42.6	174 3 US-08-463-772-50	Sequence 50, Appl
24	43.5	42.6	215 5 PCT-US93-05000-33	Sequence 33, Appl
25	43	42.4	548 4 US-10-096-571-12	Sequence 12, Appl
26	42.5	41.7	152 4 US-08-504-617-5	Sequence 5, Appl
27	42	41.2	20 4 US-09-556-877-245	Sequence 245, App

28 42 41.2 20 4 US-09-620-412C-245 Sequence 245, App
29 42 41.2 20 4 US-09-598-419-245 Sequence 245, App
30 42 41.2 542 4 US-09-489-039A-13022 Sequence 13022, A
31 41.5 40.7 237 2 US-08-464-517-24 Sequence 24, Appl
32 41.5 40.7 237 2 US-08-246-361A-24 Sequence 24, Appl
33 41.5 40.7 237 3 US-08-463-772-24 Sequence 24, Appl
34 41.5 40.7 237 5 PCT-US93-05000-24 Sequence 24, Appl
35 41 40.2 264 4 US-09-270-767-44050 Sequence 44050, A
36 41 40.2 297 4 US-09-252-991A-31076 Sequence 31076, A
37 41 40.2 397 4 US-09-491-577-70 Sequence 70, Appl
38 41 40.2 516 4 US-09-489-039A-7660 Sequence 7660, Ap
39 40.5 39.7 228 4 US-09-710-279-3190 Sequence 3190, Ap
40 40.5 39.7 263 1 US-08-276-099A-17 Sequence 17, Appl
41 40.5 39.7 263 1 US-08-781-890-17 Sequence 17, Appl
42 40.5 39.7 295 4 US-09-650-088A-10 Sequence 10, Appl
43 40.5 39.7 375 4 US-09-328-352-6191 Sequence 6191, Ap
44 40.5 39.7 502 4 US-09-710-279-3136 Sequence 3136, Ap
45 40.5 39.7 513 3 US-09-134-001C-4629 Sequence 4629, Ap

ALIGNMENTS

RESULT 1
US-09-621-976-6534
; Sequence 6534, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6534
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6534

Query Match 45.6%; Score 46.5; DB 4; Length 55;
Best Local Similarity 58.8%; Pred. No. 2.3;
Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Qy 2 MQVLLCKLLWKID-NPG 17
Db 20 WSQLLRRRLRWKDLNPG 36

RESULT 2
US-09-621-976-6991
; Sequence 6991, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6991
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6991

Query Match 45.6%; Score 46.5; DB 4; Length 55;

```

Best Local Similarity 58.8%; Pred. No. 2.3; DB 4; Length 189;
Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 2 WQVLLGKLLWKID-NPG 17
    ||:|||||
Db 20 WQVLLGKLLWKID-NPG 36
    ||:|||||

RESULT 3
US-09-543-681A-7788
; Sequence 7788, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7788
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7788

Query Match 45.6%; Score 46.5; DB 4; Length 189;
Best Local Similarity 50.0%; Pred. No. 8.9;
Matches 11; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

QY 1 DW-----QVLLGKLLWKIDNPG 17
    ||:|||||
Db 31 DWFHMPAVLLGSLPDLDPG 52
    ||:|||||

RESULT 4
US-08-464-517-6
; Sequence 6, Application US/08464517
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; FILING DATE: 16-MAY-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-464-517-6

Query Match 44.6%; Score 45.5; DB 3; Length 280;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DWQVVL-LGKLLWKI 13
    ||:|||||
Db 127 DWEVLVIGKLLWKDL 140
    ||:|||||

RESULT 5
US-08-463-772-6
; Sequence 6, Application US/08463772
; Patent No. 606501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-772-6

Query Match 44.6%; Score 45.5; DB 3; Length 280;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DWQVVL-LGKLLWKI 13
    ||:|||||

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-464-517-6

Query Match 44.6%; Score 45.5; DB 2; Length 280;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DWQVVL-LGKLLWKI 13
    ||:|||||
Db 127 DWEVLVIGKLLWKDL 140
    ||:|||||

RESULT 5
US-08-463-772-6
; Sequence 6, Application US/08463772
; Patent No. 606501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-772-6

Query Match 44.6%; Score 45.5; DB 3; Length 280;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DWQVVL-LGKLLWKI 13
    ||:|||||

```

Db 127 DWEVLVLGKIKWDL 140

RESULT 6

PCT-US93-05000-6
 ; Sequence 6, Application PC/TUS9305000
 ; GENERAL INFORMATION:
 ; APPLICANT: MITOXIX
 ; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 ; STREET: Two Militia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: US
 ; ZIP: 02173
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/05000
 ; FILING DATE: 19930525
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/888,178
 ; FILING DATE: 26-MAY-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Granahan, Patricia
 ; REGISTRATION NUMBER: 32,227
 ; REFERENCE/DOCKET NUMBER: CSHL91-02A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-861-6240
 ; TELEFAX: 616-861-9540
 ; INFORMATION FOR SEQ ID NO: 6:
 ; LENGTH: 291 amino acids
 ; SEQUENCE CHARACTERISTICS:
 ; TYPE: AMINO ACID
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; PCT-US93-05000-6

Query Match 44.6%; Score 45.5; DB 5; Length 291;
 Best Local Similarity 64.3%; Pred. No. 20;
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 DWOVL-LGKILWKI 13
 ||:|||||:
 Db 138 DWEVLVLGKIKWDL 151

RESULT 7

US-08-464-517-23
 ; Sequence 23, Application US/08464517
 ; Patent No. 5869640
 ; GENERAL INFORMATION:
 ; APPLICANT: BEACH, David H.
 ; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/464,517
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/963,308
 ; FILING DATE: 16-OCT-1992
 ; APPLICATION NUMBER: US 07/888,178
 ; FILING DATE: 26-MAY-1992
 ; PRIOR APPLICATION DATA: US 07/701,514
 ; APPLICATION NUMBER:
 ; FILING DATE: 16-MAY-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Matthew P. Vincent
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: MII-004C
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 23:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 292 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-464-517-23

Query Match 44.6%; Score 45.5; DB 2; Length 292;
 Best Local Similarity 64.3%; Pred. No. 20;
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 DWOVL-LGKILWKI 13
 ||:|||||:
 Db 139 DWEVLVLGKIKWDL 152

RESULT 8

US-08-246-361A-6
 ; Sequence 6, Application US/08246361A
 ; Patent No. 5998582
 ; GENERAL INFORMATION:
 ; APPLICANT: BEACH, David H.
 ; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/246,361A
 ; FILING DATE: 19-MAY-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/963,308
 ; FILING DATE: 16-OCT-1992
 ; PRIOR APPLICATION DATA: US 07/888,178
 ; APPLICATION NUMBER:
 ; FILING DATE: 26-MAY-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/701,514
 ; FILING DATE: 16-MAY-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Matthew P. Vincent
 ; REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: MII-004C
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-246-361A-6

Query Match 44.6%; Score 45.5; DB 2; Length 292;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DQOVL-LGKLLWKI 13
||:|||||:
Db 139 DWEVLVLGKLKNDL 152

RESULT 9
US-08-246-361A-23
Sequence 23, Application US/08246361A
Patent No. 5998582
GENERAL INFORMATION:
APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/246,361A
FILING DATE: 19-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-246-361A-23

Query Match 44.6%; Score 45.5; DB 2; Length 292;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DQOVL-LGKLLWKI 13
||:|||||:
Db 139 DWEVLVLGKLKNDL 152

QY 1 DQOVL-LGKLLWKI 13
||:|||||:
Db 139 DWEVLVLGKLKNDL 152

RESULT 10
US-08-463-772-23
Sequence 23, Application US/08463772
Patent No. 6066501
GENERAL INFORMATION:
APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,772
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-463-772-23

Query Match 44.6%; Score 45.5; DB 3; Length 292;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DQOVL-LGKLLWKI 13
||:|||||:
Db 139 DWEVLVLGKLKNDL 152

RESULT 11
PCT-US93-05000-23
Sequence 23, Application PC/TUS9305000
GENERAL INFORMATION:
APPLICANT: MITOTIX
TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive

; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US93-05000-23

Query Match 44.1%; Score 45.5; DB 5; Length 292;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 DWQVL-LGKLWKI 13
Db 139 DWEVLVLGKLKWL 152

RESULT 12
US-09-556-877-246
; Sequence 246, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probat, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 246
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-09-556-877-246

Query Match 44.1%; Score 45; DB 4; Length 20;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 GKLLWKIDNPG 17
Db 8 GKLVWKIDRLG 18

RESULT 13
US-09-556-877-247
; Sequence 247, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probat, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 247
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-09-556-877-247

Query Match 44.1%; Score 45; DB 4; Length 20;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 GKLLWKIDNPG 17
Db 3 GKLVWKIDRLG 13

RESULT 14
US-09-620-412C-246
; Sequence 246, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 246
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-09-620-412C-246

Query Match 44.1%; Score 45; DB 4; Length 20;
Best Local Similarity 72.7%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 GKLLWKIDNPG 17
Db 8 GKLVWKIDRLG 18

RESULT 15
US-09-620-412C-247
; Sequence 247, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

```

; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 247
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-09-620-412C-247

Query Match      44.1%; Score 45; DB 4; Length 20;
Best Local Similarity 72.7%; Pred. NO. 1.4;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      7 GKLLWKIDNPG 17
      |||:|||||
Db      3 GKLVWKIDRLG 13

Search completed: November 10, 2004, 14:55:36
Job time : 6.63868 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 20.8189 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-49
Perfect score: 102
Sequence: 1 DWQVLLGKLLWKIDNPGI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35325886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	100.0	18	14	US-10-092-750-49
2	52	51.0	166	15	US-10-424-599-216294
3	49	48.0	116	17	US-10-425-115-292731
4	49	48.0	144	16	US-10-767-701-56899
5	49	48.0	649	17	US-10-425-115-289398
6	48	47.1	57	16	US-10-437-963-125335
7	47.5	46.6	477	17	US-10-425-115-310459
8	47	46.1	446	16	US-10-437-963-195522
9	47	46.1	474	9	US-09-738-626-6653
10	47	46.1	652	16	US-10-437-963-130091
11	46	45.1	152	15	US-10-424-599-276306
12	46	45.1	156	15	US-10-424-599-169611
13	46	45.1	215	16	US-10-437-963-133338

RESULT 1
US-10-092-750-49
; Sequence 49, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-49

ALIGNMENTS

14	45.1	305	16	US-10-437-963-162659	Sequence 162659,
15	46	312	15	US-10-424-599-218482	Sequence 218482,
16	46	312	15	US-10-437-963-133339	Sequence 133339,
17	46	1182	16	US-10-437-963-200791	Sequence 200791,
18	46	1507	16	US-10-437-963-200793	Sequence 200793,
19	45.5	242	14	US-10-094-749-2076	Sequence 2076,
20	45.5	292	14	US-10-116-275-275	Sequence 275, App
21	45.5	292	16	US-10-408-765A-88	Sequence 88, Appl
22	45	20	9	US-09-841-132-246	Sequence 246, App
23	45	20	9	US-09-841-132-247	Sequence 247, App
24	45	197	16	US-10-767-701-32655	Sequence 32655, A
25	45	460	17	US-10-425-115-315440	Sequence 315440,
26	45	460	15	US-10-425-114-41567	Sequence 41567, A
27	45	553	9	US-09-841-132-441	Sequence 441, App
28	45	553	16	US-10-467-534-43	Sequence 43, Appl
29	45	554	15	US-10-369-435-12	Sequence 12, Appl
30	45	554	15	US-10-369-435-24	Sequence 24, Appl
31	45	556	9	US-09-841-132-398	Sequence 398, App
32	45	556	15	US-10-369-435-30	Sequence 30, Appl
33	45	556	15	US-10-312-273-47	Sequence 47, Appl
34	45	556	15	US-10-289-762-596	Sequence 596, App
35	45	686	16	US-10-437-963-130408	Sequence 130408,
36	45	712	15	US-10-425-114-57798	Sequence 57798, A
37	45	978	17	US-10-425-115-319813	Sequence 319813,
38	45	994	15	US-10-425-114-62642	Sequence 62642, A
39	45	1491	16	US-10-437-963-155569	Sequence 155569,
40	44	86	17	US-10-425-113-286181	Sequence 286181,
41	44	216	17	US-10-425-115-222672	Sequence 222672,
42	44	239	14	US-10-219-220-287	Sequence 287, App
43	44	345	9	US-09-738-626-3866	Sequence 3866, Ap
44	44	424	15	US-10-424-599-210109	Sequence 210109,
45	44	490	16	US-10-437-963-152062	Sequence 152062,

Query Match	100.0%;	Score 102;	DB 14;	Length 18;
Best Local Similarity	100.0%;	Pred. NO. 1.6e-08;		
Matches	18;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;
Qy	1	DWQVLLGKLLWKIDNPGI	18	
Db	1	DWQVLLGKLLWKIDNPGI	18	
RESULT 2				
US-10-424-599-216294				
; Sequence 216294, Application US/10424599				
; Publication No. US2004003107A1				
; GENERAL INFORMATION:				

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 216294
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_37342C.1.pap
US-10-424-599-216294

Query Match 51.0%; Score 52; DB 15; Length 166;
Best Local Similarity 50.0%; Pred. No. 6.3;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DMQVLLGKLLWKIDNP 16
|:::|:|:|:|
Db 98 DKILVDMLLWVDNP 113

RESULT 3
US-10-425-115-292731
; Sequence 292731, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 292731
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_30053C.1.pap
US-10-425-115-292731

Query Match 48.0%; Score 49; DB 17; Length 116;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DMQVLLGKLLWKIDNP 16
|:::|:|:|:|
Db 84 DKILVDMLLWVDNP 99

RESULT 4
US-10-767-701-56899
; Sequence 56899, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 56899
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30950545.pap
US-10-767-701-56899

Query Match 48.0%; Score 49; DB 16; Length 144;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DMQVLLGKLLWKIDNP 16
|:::|:|:|:|
Db 84 DKILVDMLLWVDNP 99

RESULT 5
US-10-425-115-289398
; Sequence 289398, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 289398
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(649)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_27020C.1.pap
US-10-425-115-289398

Query Match 48.0%; Score 49; DB 17; Length 649;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DMQVLLGKLLWKIDNP 16
|:::~|:|:|:|
Db 84 DKILVDMLLWVDNP 99

RESULT 6
US-10-437-963-125535
; Sequence 125535, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 125535

```
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2816C.1.pep
US-10-437-963-125535

Query Match          47.1%; Score 48; DB 16; Length 57;
Best Local Similarity 50.0%; Pred. No. 8.6;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 WQVLLGKLLWKIDNPG 17
Db 2 WKQLLGKISWKAPKSG 17

RESULT 7
US-10-425-115-310459
; Sequence 310459, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 310459
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_46196C.1.pep
US-10-425-115-310459

Query Match          46.6%; Score 47.5; DB 17; Length 477;
Best Local Similarity 30.0%; Pred. No. 90;
Matches 9; Conservative 3; Mismatches 5; Indels 13; Gaps 1;

Qy 1 DWQVLLGKLLM-----KIDNPG 17
Db 207 DWNLYLNTLFWNLWYDGSISTLSGEVDNPG 236

RESULT 8
US-10-437-963-195522
; Sequence 195522, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195522
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
```

```
; LOCATION: (1)..(446)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91462C.1.pep
US-10-437-963-195522

Query Match          46.1%; Score 47; DB 16; Length 446;
Best Local Similarity 72.7%; Pred. No. 1e+02;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LLGKLLWKIDN 15
Db 285 LLGKRLWKLEN 295

RESULT 9
US-09-738-626-6653
; Sequence 6653, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, NARATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6653
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6653

Query Match          46.1%; Score 47; DB 9; Length 474;
Best Local Similarity 64.3%; Pred. No. 1.1e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 4 VLLGKLLWKIDNPG 17
Db 15 VLLGSLWPLAAPG 28

RESULT 10
US-10-437-963-130091
; Sequence 130091, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195522
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
```

; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 130091
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_32286C.1.pap
US-10-437-963-130091

Query Match 46.1%; Score 47; DB 16; Length 652;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 DMQVLGKLLWKIDNP 16
DB 84 DKKILVDMLFWAIDNP 99

RESULT 11
US-10-424-599-276306
; Sequence 276306, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 276306
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(152)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_91523C.1.pap
US-10-424-599-276306

Query Match 45.1%; Score 46; DB 15; Length 152;
Best Local Similarity 43.8%; Pred. No. 47;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DMQVLGKLLWKIDNP 16
DB 94 DKKILVDMLFWAIDNP 109

RESULT 12
US-10-424-599-169611
; Sequence 169611, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 169611
; LENGTH: 156

; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_124172C.1.pap
US-10-424-599-169611

Query Match 45.1%; Score 46; DB 15; Length 156;
Best Local Similarity 54.5%; Pred. No. 49;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 7 GKLLWKIDNPG 17
DB 105 GKIIWEVDNNG 115

RESULT 13
US-10-437-963-133338
; Sequence 133338, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204956
; SEQ ID NO 133338
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_35218C.1.pap
US-10-437-963-133338

Query Match 45.1%; Score 46; DB 16; Length 215;
Best Local Similarity 53.8%; Pred. No. 67;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 QVLLGKLLWKIDN 15
DB 8 KALLGKWIWLEN 20

RESULT 14
US-10-437-963-162659
; Sequence 162659, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204956
; SEQ ID NO 162659
; LENGTH: 305
; TYPE: PRT

```

; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_6172C.1.pep
US-10-437-963-162659

Query Match      45.1%; Score 46; DB 16; Length 305;
Best Local Similarity 53.8%; Pred. No. 96;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      3 QVLLGKLLWKIDN 15
Db      260 KALLGKWIWKL 272

RESULT 15
US-10-424-599-218482
; Sequence 218482, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 218482
; LENGTH: 312.
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(312)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_39317C.1.pep
US-10-424-599-218482

Query Match      45.1%; Score 46; DB 15; Length 312;
Best Local Similarity 46.7%; Pred. No. 98;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      2 WQVLLGKLLWKIDNP 16
Db      70 WKKYTGVLWKIQNP 84

Search completed: November 11, 2004, 02:43:05
Job time : 21.8689 secs

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 4.41509 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-49

Perfect score: 102
Sequence: 1 DWQVLLGKLLWKIDNPGI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	50.0	309	T33259	hypothetical prote
2	50	49.0	207	G86815	NADH dehydrogenase
3	48	47.1	254	T29556	hypothetical prote
4	48	47.1	258	F97387	probable transcrip
5	48	47.1	290	AG2605	transcription regu
6	46	45.1	275	H81074	phytoene synthase-
7	45.5	44.6	292	B42822	cyclin D3 - human
8	45	44.1	354	A35829	allantoicase (EC 3
9	45	44.1	547	A32244	60K cysteine-rich
10	45	44.1	547	B43584	60K cysteine-rich
11	45	44.1	553	D71515	60K outer membr
12	45	44.1	554	C81671	60K cysteine-rich
13	45	44.1	556	S12602	60 kDa Cysteine-ri
14	45	44.1	556	A86560	hypothetical prote
15	45	44.1	673	T48012	hypothetical prote
16	44	43.1	107	C85823	hypothetical prote
17	44	43.1	107	E90976	hypothetical prote
18	44	43.1	210	C70627	hypothetical prote
19	44	43.1	545	T36123	probable lysyl-trn
20	44	43.1	547	B64963	membrane protein y
21	44	43.1	1025	T31014	hypothetical prote
22	43.5	42.6	135	A82373	mannose-6-phosphat
23	43.5	42.6	432	D64743	cell cycle protein
24	43	42.2	276	D81968	probable transfe
25	43	42.2	279	T48013	hypothetical prote
26	43	42.2	355	F75050	GTP-binding protei
27	43	42.2	778	T39047	hypothetical prote
28	42.5	41.7	896	F96523	hypothetical prote
29	42	41.2	300	F83804	cation efflux syst

30 42 41.2 306 1 S18215
31 42 41.2 402 2 G75011
32 42 41.2 512 2 T13635
33 42 41.2 548 2 G86062
34 42 41.2 548 2 F91216
35 42 41.2 579 2 JC7170
36 42 41.2 736 2 T40510
37 42 41.2 810 2 A86222
38 42 41.2 865 2 AC2236
39 42 41.2 1660 2 A84647
40 41.5 40.7 237 2 C40035
41 41.5 40.7 293 2 JC4012
42 41.5 40.7 553 2 AB1475
43 41.5 40.7 639 2 T24972
44 41.5 40.7 698 2 T24970
45 41 40.2 240 2 S75753

ALIGNMENTS

RESULT 1

T33259
hypothetical protein C24B9.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33259
R;Murray, J.; Wohlmann, P.; Langston, Y.; O'Neal, D.
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of C. elegans cosmid C24B9.
A;Reference number: Z21310
A;Accession: T33259
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-309 <MUR>
A;Cross-references: UNIPROT:O76432; EMBL:AF068709; PIDN:AAC19253.1; GSPDB:GN000023; CESI:
A;Experimental source: strain Bristol N2; clone C24B9
C;Genetics:
A;Gene: CESP:C24B9.2

A;Map position: 5
C;Superfamily: Caenorhabditis elegans hypothetical protein C16D9.7

Query Match 50.0%; Score 51; DB 2; Length 309;
Best Local Similarity 40.0%; Pred. No. 2.2;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Cy 1 DWQVLLGKLLWKIDN 15
Db 258 EMMIVLAQLWQLDH 272

RESULT 2

G86815
NADH dehydrogenase yphA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: G86815
R;Bolorin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrh
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: G86815
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-207 <STO>
A;Cross-references: UNIPROT:QSCFF1; GB:AE005176; PID:g12724526; PIDN:AAK05625.1; GSPDB:
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yphA

Query Match 49.0%; Score 50; DB 2; Length 207;
Best Local Similarity 61.1%; Pred. No. 2;

Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDNPGI 18
 DB 57 DWQELLQKSDWVIDAVGI 74

RESULT 3
 T29556
 hypothetical protein C16D9.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T29556
 R:Gattung, S.; Le, T.T.
 submitted to the EMBL Data Library, July 1996
 A:Description: The sequence of C. elegans cosmid C16D9.
 A:Reference number: Z20640
 A:Accession: T29556
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-254 <CAT>
 A:Cross-references: UNIPROT:Q22901; EMBL:U64858; PIDN:AAB18286.1; GSPDB:GN000023; CESP:CI
 A:Experimental source: strain Bristol N2; clone C16D9
 C:Genetics:
 A:Gene: CESP:C16D9.7
 A:Map position: 5
 A:Introns: 123/3; 166/3; 213/1; 238/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein C16D9.7

Query Match 47.1%; Score 48; DB 2; Length 254;
 Best Local Similarity 40.0%; Pred. No. 5.2; Mismatches 2; Indels 0; Gaps 0;
 Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDN 15
 DB 198 EMIIVLAQLLWQLDH 212

RESULT 4
 F97387
 probable transcription regulator (P24489) [imported] - Agrobacterium tumefaciens (strain
 C:Species: Agrobacterium tumefaciens
 C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C:Accession: F97387
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Curolo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: F97387
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-258 <KOR>
 A:Cross-references: UNIPROT:Q8UIQ4; GB:AE007869; PIDN:AAK86055.1; PID:gl5155128; GSPDB:G
 C:Genetics:
 A:Gene: AGR_C 407
 A:Map position: circular chromosome
 C:Superfamily: arabinose operon regulatory protein

Query Match 47.1%; Score 48; DB 2; Length 258;
 Best Local Similarity 60.0%; Pred. No. 5.3; Mismatches 6; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDN 15
 DB 129 DWQSRLLSDDLWKHSN 143

RESULT 5
 AG2605
 transcription regulator, AraC family Atu0239 [imported] - Agrobacterium tumefaciens (str
 C:Species: Agrobacterium tumefaciens
 C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C:Accession: AG2605
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
 i Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AG2605
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-290 <KOR>
 A:Cross-references: UNIPROT:Q8UIQ4; GB:AE008688; PIDN:AAL41261.1; PID:gl7738567; GSPDB:
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu0239
 A:Map position: circular chromosome

Query Match 47.1%; Score 48; DB 2; Length 290;
 Best Local Similarity 60.0%; Pred. No. 6; Mismatches 6; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDN 15
 DB 161 DWQSRLLSDDLWKHSN 175

RESULT 6
 H81074
 Phycocyanin synthase-related protein NMB1521 [imported] - Neisseria meningitidis (strain M
 C:Species: Neisseria meningitidis
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C:Accession: H81074
 R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, C.; Fleischmann, R.D.; Dougherty, B.A.
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
 A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: AB1000; MUID:20175755; PMID:10710307
 A:Accession: H81074
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-275 <TET>
 A:Cross-references: UNIPROT:Q8JYL8; GB:AE002501; GB:AE002098; NID:g7226755; PIDN:AAF418
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB1521

Query Match 45.1%; Score 46; DB 2; Length 275;
 Best Local Similarity 58.3%; Pred. No. 12; Mismatches 2; Indels 0; Gaps 0;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWK 12
 DB 264 DWLMLKLLWK 275

RESULT 7
 B42822
 Cyclin D3 - human
 N:Alternate names: PRAD1/cyclin D1 proto-oncogene homolog
 C:Species: Homo sapiens (man)
 C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 12-Jul-2004
 C:Accession: B42822; A44022; I60850
 R:Xiong, Y.; Menninger, J.; Beach, D.; Ward, D.C.
 Genomics 13, 575-584, 1992
 A>Title: Molecular cloning and chromosomal mapping of CCND genes encoding human D-type
 A:Reference number: A42822; MUID:92347851; PMID:1386336
 A:Accession: B42822
 A:Molecule type: mRNA
 A:Residues: 1-292 <XIO>

A;Cross-references: UNIPROT:P30281; GB:M90814; NID:G180002; PIDN:AAA51927.1; PID:G180003
 A;Note: sequence extracted from NCBI backbone (NCBIN:109687, NCBI:P:109688)
 B;Morokura, T.; Keyonaris, K.; Kronenberg, H.M.; Arnold, A.
 J. Biol. Chem. 267, 20412-20415, 1992
 A;Title: Cloning and characterization of human cyclin D3, a cDNA closely related in sequence to cyclin D1
 A;Reference number: A44022; MUID:93015922; PMID:1383201
 A;Accession: A44022
 A;Molecule type: mRNA
 A;Residues: 1-258; S', 260-292 <NOT>
 A;Cross-references: GB:M92287; NID:G181246; PIDN:AAA52137.1; PID:G181247
 A;Experimental source: placenta
 A;Note: sequence extracted from NCBI backbone (NCBI:P:115738)
 R;Inaba, T.; Matsushima, H.; Valentine, M.; Rousset, M.F.; Sherr, C.J.; Look, A.T.
 Genomics 13, 565-574, 1992
 A;Title: Genomic organization, chromosomal localization, and independent expression of human cyclin D3
 A;Reference number: A42821; MUID:92347850; PMID:1386335
 A;Accession: I60850
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 52-237 <RES>
 A;Cross-references: GB:M98087; NID:G180014; PIDN:AAA51929.1; PID:G180016
 C;Genetics:
 A;Gene: GDB:CCND3
 A;Cross-references: GDB:128969; OMIM:123834
 A;Map position: 6p21-6p21
 A;Introns: 66/3; 138/3; 192/1
 A;Note: introns line may be incomplete
 C;Superfamily: cyclin, A/B/D/E type
 C;Keywords: cell cycle control

Query Match 44.6%; Score 45.5; DB 2; Length 292;
 Best Local Similarity 64.3%; Pred. No. 15;
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 DWQVL-LGKLLWKI 13
 |||||
 DB 139 DWEVLVLCXKWL 152

RESULT 8
 A35829
 allantoicase (EC 3.5.3.4) - Neurospora crassa
 C;Species: Neurospora crassa
 C;Date: 30-Nov-1990 #sequence_revision 13-Sep-1991 #text_change 09-Jul-2004
 C;Accession: A35829
 R;Lee, H.; Fu, Y.H.; Marzluf, G.A.
 Biochemistry 29, 8779-8787, 1990
 A;Title: Nucleotide sequence and DNA recognition elements of alc, the structural gene for allantoicase
 A;Reference number: A35829; MUID:91104776; PMID:2148685
 A;Accession: A35829
 A;Molecule type: mRNA
 A;Residues: 1-354 <LER>
 A;Cross-references: UNIPROT:P18407; GB:J02927; NID:G168748; PIDN:AAA33556.1; PID:G168749
 A;Note: the authors translated the codon TCC for residue 42 as Leu, CTG for residue 140
 C;Superfamily: Allantoicase
 C;Keywords: hydrolase

Query Match 44.1%; Score 45; DB 2; Length 354;
 Best Local Similarity 30.4%; Pred. No. 22;
 Matches 7; Conservative 5; Mismatches 3; Indels 8; Gaps 1;

QY 2 WQVLIG-----KLLWKIDNP 16
 :||| :|||
 DB 143 WETILGVCEGSGRQFCWKLENP 165

RESULT 9
 A32244
 60K cysteine-rich outer membrane protein 2 precursor, serotype L1 and L2 - Chlamydia trachomatis
 C;Species: Chlamydia trachomatis
 C;Date: 12-Oct-1989 #sequence_revision 27-Jun-1994 #text_change 09-Jul-2004
 C;Accession: A32244; A43584; A36043; A30472; J0419; S18981; S24277

R;Allen, J.E.; Stephens, R.S.
 J. Bacteriol. 171, 285-291, 1989
 A;Title: Identification by sequence analysis of two-site posttranslational processing of the 60K cysteine-rich outer membrane protein 2 precursor, serotype E - Chlamydia trachomatis
 A;Reference number: A32244; MUID:89123030; PMID:2914847
 A;Accession: A32244
 A;Molecule type: DNA
 A;Residues: 1-547 <ALL>
 A;Cross-references: UNIPROT:P21354; GB:M23001; NID:G144552; PIDN:AAA23152.1; PID:G144553
 A;Experimental source: strain L2/434/Bu
 A;Note: parts of this sequence, including the amino ends of the precursor and mature protein, are identical to those of the 60K cysteine-rich outer membrane protein 2 precursor, serotype E - Chlamydia trachomatis
 R;de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
 Infect. Immun. 59, 1196-1201, 1991
 A;Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton protein of Chlamydia trachomatis
 A;Reference number: A43584; MUID:91147205; PMID:1997423
 A;Accession: A43584
 A;Molecule type: DNA
 A;Residues: 1-547
 A;Cross-references: GB:M23001; NID:G144552; PIDN:AAA23152.1; PID:G144553
 A;Experimental source: serovar 2, strain L2/434/Bu
 R;Wahlberg, J.; Lundberg, J.; Hultman, T.; Uhlen, M.
 Proc. Natl. Acad. Sci. U.S.A. 87, 6569-6573, 1990
 A;Title: General colorimetric method for DNA diagnostics allowing direct solid-phase PCR
 A;Reference number: A36043; MUID:90370827; PMID:2118652
 A;Accession: A36043
 A;Molecule type: DNA
 A;Residues: 294-402 <WAH>
 A;Experimental source: serotype L2
 R;Lambden, P.R.; Everson, J.S.; Ward, M.E.; Clarke, I.N.
 Gene 87, 105-112, 1990
 A;Title: Sulfur-rich proteins of Chlamydia trachomatis: developmentally regulated transmembrane proteins
 A;Reference number: JQ0514; MUID:90236284; PMID:2332164
 A;Accession: A30472
 A;Molecule type: DNA
 A;Residues: 1-46; 528-547 <LAM>
 A;Cross-references: GB:M35148; GB:M23180; GB:M35161; NID:G144485
 A;Experimental source: serotype L1
 R;Clarke, I.N.; Ward, M.E.; Lambden, P.R.
 Gene 71, 307-314, 1988
 A;Title: Molecular cloning and sequence analysis of a developmentally regulated cysteine-rich outer membrane protein of Chlamydia trachomatis
 A;Reference number: J0419; MUID:89138006; PMID:3066701
 A;Accession: J0419
 A;Molecule type: DNA
 A;Residues: 30-547 <CLA>
 A;Cross-references: GB:M35148; NID:G144485; PIDN:AAA23119.1; PID:G144487
 A;Experimental source: serotype L1
 C;Genetics:
 A;Gene: omp2; ompB
 C;Function:
 A;Description: associated with differentiation of reticulate bodies into elementary bodies
 A;Note: essential for the structural integrity of the outer envelope of the elementary bodies
 C;Superfamily: 60K cysteine-rich outer membrane protein
 C;Keywords: membrane protein; virulence
 F;1-22/Domain: signal sequence #status predicted <SIG>
 F;23-40/Domain: propeptide #status experimental <PRO>
 F;41-547/Product: 60K cysteine-rich outer membrane protein 2 #status experimental <MAT>

Query Match 44.1%; Score 45; DB 1; Length 547;
 Best Local Similarity 72.7%; Pred. No. 36;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GKLLWKIDNP 17
 |||||
 DB 159 GKLLWKIDRLG 169

RESULT 10
 B43584
 60K cysteine-rich outer membrane protein 2 precursor, serotype E - Chlamydia trachomatis
 C;Species: Chlamydia trachomatis
 C;Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
 C;Accession: B43584; S13120; S18979; S24275
 R;de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
 Infect. Immun. 59, 1196-1201, 1991

A;Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton
A;Reference number: A43584; MUID:91147205; PMID:1997423
A;Accession: B43584
A;Molecule type: DNA
A;Residues: 1-547
A;Cross-references: UNIPROT:P23603; EMBL:X54389; NID:G40763; PIDN:CAA38259.1; PID:G40764
A;Experimental source: strain Bour, serotype E
R;Colles, A.M.; Allan, I.; Pearce, J.H.
Nucleic Acids Res. 18, 6713, 1990
A;Title: The nucleotide and derived amino acid sequence of the omp2 gene of Chlamydia tr
A;Reference number: S13120; MUID:91067486; PMID:2251143
A;Accession: S13120
A;Molecule type: DNA
A;Residues: 1-32, 'FT', 35-120, 'L', 122-131, 'A', 133-457, 'S', 459-547 <COL>
A;Cross-references: EMBL:X55903; NID:G40724; PIDN:CAA39396.1; PID:G40725
A;Experimental source: strain DK20, serotype E
C;Genetics:
A;Gene: omp2; omcB
C;Function:
A;Description: associated with differentiation of reticulate bodies into elementary bodi
A;Note: essential for the structural integrity of the outer envelope of the elementary b
C;Superfamily: 60K cysteine-rich outer membrane protein
C;Keywords: membrane protein; virulence
F;1-2/Domain: signal sequence #status predicted <SIG>
F;23-40/Domain: propeptide #status predicted <PRO>
F;41-547/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>
Query Match 44.1%; Score 45; DB 2; Length 547;
Best Local Similarity 72.7%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 GKLLWKIDNPG 17
|||:|||||
DB 159 GKLVWKIDRLG 169
|||:|||||
RESULT 11
D71515
60K cysteine-rich outer membrane protein 2 precursor serotypes B, C, and D - Chlamydia b
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence revision 13-Sep-1998 #text_change 09-Jul-2004
C;Accession: D71515; C43584; S11673
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A;Reference number: A71570; MUID:99000809; PMID:9784136
A;Accession: D71515
A;Molecule type: DNA
A;Residues: 1-553 <ARN>
A;Cross-references: UNIPROT:P26758; UNIPROT:P18151; GB:AE001317; GB:AE001273; NID:G43288
A;Experimental source: serotype D, strain UW-3/Cx
R;de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Markoff, B.A.; Peterson, E.M.
Infect. Immun. 59, 1195-1201, 1991
A;Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodalton
A;Reference number: A43584; MUID:91147205; PMID:1997423
A;Accession: C43584
A;Molecule type: DNA
A;Residues: 7-238, 'V', 240-553
A;Cross-references: GB:X54388; NID:G40760; PIDN:CAA38257.1; PID:G40761
A;Experimental source: serotype C
R;Watson, M.W.; Lambden, P.R.; Ward, M.E.; Clarke, I.N.
FEMS Microbiol. Lett. 65, 293-297, 1989
A;Title: Chlamydia trachomatis 60 kDa cysteine rich outer membrane protein: sequence hom
A;Reference number: S11673
A;Accession: S11673
A;Molecule type: DNA
A;Residues: 7-553 <MAT>
A;Cross-references: EMBL:X53510; NID:G40681; PIDN:CAA37588.1; PID:G40683
A;Experimental source: serotype B
C;Genetics:
A;Gene: omp2; omcB
C;Function:
A;Description: associated with differentiation of reticulate bodies into elementary bodi

A;Note: essential for the structural integrity of the outer envelope of the elementary
C;Superfamily: 60K cysteine-rich outer membrane protein
C;Keywords: membrane protein; virulence
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-46/Domain: propeptide #status predicted <PRO>
F;47-553/Product: 60K cysteine-rich outer membrane protein 2 #status predicted <MAT>
Query Match 44.1%; Score 45; DB 2; Length 553;
Best Local Similarity 72.7%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 GKLLWKIDNPG 17
|||:|||||
DB 165 GKLVWKIDRLG 175
|||:|||||
RESULT 12
C81671
60 kDa outer membrane protein TC0727 [imported] - Chlamydia muridarum (strain Nigg)
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: C81671
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Nelson, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: C81671
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-554 <TET>
A;Cross-references: UNIPROT:Q9PJV0; GB:AE002341; GB:AE002160; NID:G7190754; PIDN:AAF395
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0727
C;Superfamily: 60K cysteine-rich outer membrane protein
Query Match 44.1%; Score 45; DB 2; Length 554;
Best Local Similarity 72.7%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 7 GKLLWKIDNPG 17
|||:|||||
DB 166 GKLVWKIDRLG 176
|||:|||||
RESULT 13
S12602
60K cysteine-rich outer membrane protein precursor [similarity] - Chlamydia pneumoniae
C;Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C;Date: 30-Sep-1993 #sequence_revision 27-Jun-1994 #text_change 09-Jul-2004
C;Accession: S12602; H72063; A81604
R;Watson, M.W.; Al-Mahdawi, S.; Lamden, P.R.; Clarke, I.N.
Nucleic Acids Res. 18, 5299, 1990
A;Title: The nucleotide sequence of the 60kDa cysteine rich outer membrane protein of C
A;Reference number: S12602; MUID:90384850; PMID:2402463
A;Accession: S12602
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-556 <WAT>
A;Cross-references: UNIPROT:P23700; GB:X53511; NID:G550564; PIDN:CAA37590.1; PID:G55056
A;Experimental source: isolate IOL-207
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: H72063
A;Molecule type: DNA
A;Residues: 1-556 <ARN>
A;Cross-references: GB:AE001640; GB:AE001363; NID:G4376845; PIDN:AD18697.1; PID:G43768
A;Experimental source: strain CWJ029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg

Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
 A:Reference number: A81500; PMID:20150255; PMID:10684935
 A:Accession: A81504
 A:Molecule type: DNA
 A:Residues: 1-556 <REA>
 A:Cross-references: GB:AE002180; GB:AE002161; NID:g7189117; PIDN:AAF38068.1; PID:g718912
 A:Experimental source: strain AR39, HL cells
 C:Comment: This protein is associated with the differentiation of reticulate bodies into
 . It may also be an important virulence factor.
 C:Genetics:
 A:Gene: omcB; CF0195
 C:Superfamily: 60K cysteine-rich outer membrane protein
 C:Keywords: membrane protein; virulence
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-40/Domain: propeptide #status predicted <PRO>
 F:41-556/Product: 60K cysteine-rich outer membrane protein #status predicted <MAT>

Query Match 44.1%; Score 45; DB 1; Length 556;
 Best Local Similarity 72.7%; Pred. No. 36;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GKLLWKIDNPG 17
 |||:|||||
 DB 168 GKLVWKIDRLG 178

RESULT 14
 A86560
 60 kDa Cysteine-rich omp [imported] - Chlamydia pneumoniae (strain J138)
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: A86560
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; PMID:20330349; PMID:10871362
 A:Accession: A86560
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-556 <STO>
 A:Cross-references: UNIPROT:P23700; GB:BA000008; NID:g8978928; PIDN:BA98763.1; GSPDB:GN
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: omcB
 C:Superfamily: 60K cysteine-rich outer membrane protein

Query Match 44.1%; Score 45; DB 2; Length 556;
 Best Local Similarity 72.7%; Pred. No. 36;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 GKLLWKIDNPG 17
 |||:|||||
 DB 168 GKLVWKIDRLG 178

RESULT 15
 T48012
 Hypothetical protein T17J13.160 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
 C:Accession: T48012
 R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Mayer,
 submitted to the Protein Sequence Database, February 2000
 A:Reference number: Z24482
 A:Accession: T48012
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-673 <RIE>
 A:Cross-references: UNIPROT:Q9MIQ4; EMBL:AL138651
 A:Experimental source: cultivar Columbia; BAC clone T17J13
 C:Genetics:
 A:Map position: 3

A:Introns: 99/1; 605/2
 A:Note: T17J13.160
 Query Match 44.1%; Score 45; DB 2; Length 673;
 Best Local Similarity 43.8%; Pred. No. 45;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DWQVLLGKLLWKIDNP 16
 |||:|||||
 DB 105 DKILVDMFLFWALDNP 120

Search completed: November 10, 2004, 14:52:18
 Job time : 5.41509 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 23.9774 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-49
Perfect score: 102
Sequence: 1 DWQVLGKLLWKIDNPGI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt.02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	50.0	309	2	O76432 caenorhabdi
2	51	50.0	996	2	O9FNP6 arabidopsis
3	50	49.0	207	2	O8CFF1 lactococcus
4	49	48.0	236	2	O9TYH3 schistosoma
5	49	48.0	488	2	O9TYH4 schistosoma
6	48	47.1	257	2	O22901 caenorhabdi
7	48	47.1	258	2	O7D1V9 agrobacteri
8	48	47.1	290	2	O8UIQ4 acrobacteri
9	48	47.1	352	2	O6BW53 debaryomyce
10	48	47.1	387	2	O8PDP2 xanthomonas
11	48	47.1	389	2	O8PQL6 xanthomonas
12	48	47.1	1864	2	O95Z87 leishmania
13	47	46.1	296	2	O89ZE9 bacteroides
14	47	46.1	474	2	O8NLS8 corynebacte
15	47	46.1	474	2	CAF20884 corynebac
16	47	46.1	552	2	O7XIX3 oryza sativ
17	47	46.1	1863	2	O9H0J4 homo sapien
18	47	46.1	2314	2	O6BFH5 paramecium
19	46	45.1	178	2	O46163 chlamydia
20	46	45.1	178	2	O46165 chlamydia
21	46	45.1	212	2	O7VT35 bordetella
22	46	45.1	212	2	O7WFI2 bordetella
23	46	45.1	267	1	ADHC MOUSE
24	46	45.1	267	2	BAC38101 mus muscu
25	46	45.1	295	2	O9UYL8 neisseria m
26	46	45.1	395	2	O9X411 chlamydia
27	46	45.1	426	2	O8NEF8 homo sapien
28	46	45.1	472	2	O6ZMY6
29	46	45.1	472	2	BAD18589
30	46	45.1	481	2	O7UJPS rhodospirill
31	46	45.1	534	2	O9ZAW0 chlamydia

32 46 45.1 658 2 O6K9U1 oryza sativ
33 46 45.1 835 2 O7NN31 gloeobacter
34 46 45.1 1948 2 O7PUO9 anopheles g
35 45.5 44.6 292 1 CGD3 HUMAN
36 45 44.1 85 2 O02664 podospora a
37 45 44.1 173 2 O9AF82 chlamydia s
38 45 44.1 176 2 O9AG10 chlamydia p
39 45 44.1 183 2 O6LCH5 chlamydia p
40 45 44.1 183 2 O46164 chlamydia p
41 45 44.1 183 2 AAB00573 chlamydia
42 45 44.1 363 2 O6FNU2 candida gla
43 45 44.1 395 2 O9S6B3 chlamydia p
44 45 44.1 410 2 O96LM3 homo sapien
45 45 44.1 500 2 O6WGP4 chlamydia s

ALIGNMENTS

RESULT 1
O76432 PRELIMINARY; PRT; 309 AA.
AC O76432;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Hypothetical protein C24B9.2.
GN Name=C24B9.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Murray J., Wohldmann P., Langston Y., O'Neal D.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF068709; AAC19253.1; -.
DR PIR; T33259; T33259.
DR Wormpep; C24B9.2; CEI1744.
DR InterPro; IPR002651; DUF32.
DR Pfam; PF01748; DUF32; 1.
KW Hypothetical protein.
SQ SEQUENCE 309 AA; 35650 MW; 7AC2FD1FB1B3C27C CRC64;

```

Query Match          50.0%; Score 51; DB 2; Length 309;
Best Local Similarity 40.0%; Pred. No. 12;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDN 15
DB 258 EWMIVLAQLLWQLDH 272

RESULT 2
Q9FNP6 PRELIMINARY; PRT; 996 AA.
ID Q9FNP6
AC Q9FNP6;
DT 01-JUN-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2004 (TrEMBLrel. 16, Last sequence update)
DE Emb|CAB71880.1;
OS Arabidopsis thaliana (Mouse-ear cress);
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98069011; PubMed=9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned P1 clones.";
RL DNA Res. 4:291-300(1997).
DR EMBL; AB006696; BAB10381.1; --
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007491; DUF537.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR003604; Znf_U1.
DR Pfam; PF04396; DUF537; 1.
DR Pfam; PF00096; zf-C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 5.
DR SMART; SM00451; Znf_U1; 4.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
SQ SEQUENCE 996 AA; 111222 MW; C824826A49AD7515 CRC64;

Query Match          50.0%; Score 51; DB 2; Length 996;
Best Local Similarity 56.2%; Pred. No. 42;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDN 16
DB 88 DKILVDMLLWALDNP 103

RESULT 3
Q9CFF1 PRELIMINARY; PRT; 207 AA.
ID Q9CFF1
AC Q9CFF1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NADH dehydrogenase.
GN Names=ypha; OrderedLocustNames=L11527;
OC Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauder S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus

```

```

RT lactis esp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
DR EMBL; AB006383; AAK03625.1; --
DR PIR; G86815; G86815.
KW Complete proteome.
SQ SEQUENCE 207 AA; 23410 MW; B4BEO3A872746A3 CRC64;

Query Match          49.0%; Score 50; DB 2; Length 207;
Best Local Similarity 61.1%; Pred. No. 12;
Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDNPGI 18
DB 57 DWQELLQKSCWVIDAVGI 74

RESULT 4
Q9TYH3 PRELIMINARY; PRT; 236 AA.
ID Q9TYH3
AC Q9TYH3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Serine protease SmSPI light chain (Fragment).
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Porto Rico;
RA Cocude C.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AJ011562; CAA09692.1; --
DR HSP; P00766; 1CHG.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR01254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1.
DR InterPro; IPR005003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON TER
SQ SEQUENCE 236 AA; 26224 MW; FF9C3CDFBAC2BA4 CRC64;

Query Match          48.0%; Score 49; DB 2; Length 236;
Best Local Similarity 47.1%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 WQVLLGKLLWKIDNPGI 18
DB 205 WNCIMGKRLCKTRNPGV 221

RESULT 5
Q9TYH4 PRELIMINARY; PRT; 488 AA.
ID Q9TYH4
AC Q9TYH4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Serine protease SmSPI (Fragment).
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.

```



```

OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=porto Rico;
RA Cocude C.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AJ011561; CAA09691.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004235; F:trypsin activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00057; Ldl_recept_a; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00192; LDLa; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01180; CUE; 1.
DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS00068; LDLRA_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
SQ SEQUENCE 488 AA; 59209 MW; 29F26792C2014FA6 CRC64;

Query Match 48.0%; Score 49; DB 2; Length 488;
Best Local Similarity 47.1%; Pred. No. 41;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 WQVLLGKLLWKIDNPGI 18
DB 457 WNCINGKELCKTRNPGV 473

RESULT 6
ID Q22901 PRELIMINARY; PRT; 257 AA.
AC Q22901;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C16D9.7.
CN ORFNames=C16D9.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gattung S.; Le T.T.;
RA Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gattung S.; Le T.T.;
RA Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;

```

```

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; U64858; AAB18286.2; -.
DR PIR; T29556; T29556.
DR WormPep; C16D9.7; CE06843.
DR InterPro; IPR002651; DUF32.
DR Pfam; PF01748; DUF32; 1.
KW Hypothetical protein.
SQ SEQUENCE 257 AA; 29578 MW; B2784689FA6664CE CRC64;

Query Match 47.1%; Score 48; DB 2; Length 257;
Best Local Similarity 40.0%; Pred. No. 31;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDN 15
DB 198 EMIIVLAQLLWQLDH 212

RESULT 7
ID Q7D1V9 PRELIMINARY; PRT; 258 AA.
AC Q7D1V9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGR_C_407P.
CN OrderedLocusNames=AGR_C_407;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cerson;
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RA "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
CC -!- SIMILARITY: Contains 1 HTH arnC/xyIs-type DNA-binding domain.
DR EMBL; AS007963; AAK86055.1; -.
DR InterPro; IPR003313; AracC_binding.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR000005; HTHAracC.
DR Pfam; PF02311; AracC_binding; 1.
DR Pfam; PF00165; HTH_Arac; 1.
DR PRINTS; PR00032; HTHARAC.
DR SMART; SM00342; HTH_ARAC; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW DNA-binding; Transcription_regulation.
SQ SEQUENCE 258 AA; 29109 MW; 3D6D0256B4D65225 CRC64;

Query Match 47.1%; Score 48; DB 2; Length 258;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 DWQVLLGKLLWKIDN 15
DB 129 DWQSRSLDILLWXHN 143

RESULT 8
Q8UIQ4

```

```

ID Q8UIQ4 PRELIMINARY; PRT; 290 AA.
AC Q8UIQ4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Descriptive text, AraC family.
DE Transcriptional regulator, AraC family.
GN OrderedLocusNames=Atu0239;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=174299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
CC -!- SIMILARITY: Contains 1 HTH arac/xyis-type DNA-binding domain.
DR EMBL; AE008996; AAL41261.1; --
DR PIR; AG2605; AG2605.
DR PIR; F97387; F97387.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003313; AraC binding.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR000005; HTHARAC.
DR Pfam; PF02311; AraC binding; 1.
DR Pfam; PF00165; HTH_AraC; 1.
DR PRINTS; PR00032; HTHARAC.
DR SMART; SM00342; HTH_ARAC; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW Complete proteome; DNA-binding; Transcription regulation.
SQ SEQUENCE 290 AA; 32633 MW; C6F3643FA611B58A CRC64;

Query Match 47.1%; Score 48; DB 2; Length 290;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 DWQVLGKLLWKIDN 15
DB 161 DWQSLSDLLWKHSN 175

RESULT 9
Q6BW53 PRELIMINARY; PRT; 352 AA.
ID Q6BW53;
AC Q6BW53;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Chromosome B of strain CBS767 of Debaryomyces Hansenii.
GN ORFNames=DEBA0B143779;
OS Debaryomyces Hansenii (Yeast) (Torulasporea hansenii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=4959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RX GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,

```

```

RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Gropi A.,
RA Hantraye F., Hennequin C., Jaumiaux N., Joyet P., Kachouri R.,
RA Karrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellens S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Zeniène D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Sniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RA Genoscope; (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2004); CAG85577.1; --
DR EMBL; CR382134; CAG85577.1; --
SQ SEQUENCE 352 AA; 40083 MW; 40505E557CC2F7EC CRC64;

Query Match 47.1%; Score 48; DB 2; Length 352;
Best Local Similarity 35.3%; Pred. No. 42;
Matches 6; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 WQVLGKLLWKIDNPGI 18
DB 315 WHIRLSIFWEIDNQDV 331

RESULT 10
Q8PDF2 PRELIMINARY; PRT; 387 AA.
ID Q8PDF2;
AC Q8PDF2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nucleoside hydrolase.
GN OrderedLocusNames=XCC0293;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Pereira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezra R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
RL EMBL; AE012125; AAM39612.1; --
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR001910; I/U_nucleid_hydro.
DR Pfam; PF01156; I/U_nuc_hydro; 1.
DR ProDom; PD007736; I/U_nucleid_hydro; 1.
KW Complete proteome; Hydrolase.

```

```

SQ SEQUENCE 387 AA; 42450 MW; 44684P6D4093D4D2 CRC64;
Query Match 47.1%; Score 48; DB 2; Length 387;
Best Local Similarity 63.6%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 WQVLIGKLLWK 12
   : : : : :
Db 88 WEALYGLVWK 98

RESULT 11
Q8PQL6 PRELIMINARY; PRT; 389 AA.
AC Q8PQL6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nucleoside hydrolase.
GN OrderedLocusNames=XAC0307;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Fatah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AB011656; AAM35199.1; -.
GO GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR001910; I/U nucleic hydro.
DR Pfam; PF01156; I/U nuc hydro; 1.
DR ProDom; PD007736; I/U nucleic hydro; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 389 AA; 42765 MW; BC69BCDA1BA55D97 CRC64;
Query Match 47.1%; Score 48; DB 2; Length 389;
Best Local Similarity 63.6%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 WQVLIGKLLWK 12
   : : : : :
Db 106 WEALYGLVWK 116

RESULT 12
Q95Z87 PRELIMINARY; PRT; 1864 AA.
AC Q95Z87;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-NAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein LT.09.
GN Name=LT.09; Synonyms=P131.01;
OS Leishmania major.

```

```

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.P.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL389894; CAC44916.1; -.
DR InterPro; IPR002110; ANK.
DR PRINTS; PR01415; ANKYRIN.
KW Hypothetical protein.
SQ SEQUENCE 1864 AA; 201407 MW; C999DB258BD104E3 CRC64;
Query Match 47.1%; Score 48; DB 2; Length 1864;
Best Local Similarity 43.8%; Pred. No. 24e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 WQVLIGKLLWKIDNPG 17
   : : : : :
Db 1815 WQLLAGAKVWRFPMDPG 1830

RESULT 13
Q89ZE9 PRELIMINARY; PRT; 296 AA.
AC Q89ZE9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BT4428;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AB016945; AAO79533.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 296 AA; 32943 MW; 6BF99EBB76F9FA67 CRC64;
Query Match 46.1%; Score 47; DB 2; Length 296;
Best Local Similarity 63.6%; Pred. No. 51;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 7 GKLLWKIDNPG 17
   : : : : :
Db 168 GKVVWKLDVPG 178

RESULT 14
Q8NLS8 PRELIMINARY; PRT; 474 AA.
ID Q8NLS8
AC Q8NLS8; Q6M127;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical membrane protein (Putative membrane protein).

```

GN OrderedLocusNames=Cg12859, Cg3165;
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RX PubMed=12948626;
 RA Kallinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
 RA Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
 RA Goessmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
 RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
 RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegraebe I.,
 RA Tauch A.;
 RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
 RT and its impact on the production of L-aspartate-derived amino acids
 RT and vitamins.";
 RL J. Biotechnol. 104:5-25(2003).
 DR EMBL; AP005283; BAC00253.1; --
 DR EMBL; BX927156; CAF20884.1; --
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 474 AA; 50594 MW; 3810825D5298DDB3 CRC64;

Query Match 46.1%; Score 47; DB 2; Length 474;
 Best Local Similarity 64.3%; Pred. No. 84;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 VLLGKLLWKIDNPG 17
 DB 15 VLLGSLWPLAAPG 28

RESULT 15

CAF20884 PRELIMINARY; PRT; 474 AA.
 AC CAF20884;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative membrane protein.
 GN Cg3165.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RX PubMed=12948626;
 RA Kallinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
 RA Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
 RA Goessmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
 RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
 RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegraebe I.,
 RA Tauch A.;
 RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
 RT and its impact on the production of L-aspartate-derived amino acids
 RT and vitamins.";
 RL J. Biotechnol. 104:5-25(2003).
 DR EMBL; BX927156; CAF20884.1; --
 SQ SEQUENCE 474 AA; 50594 MW; 3810825D5298DDB3 CRC64;

Query Match 46.1%; Score 47; DB 2; Length 474;
 Best Local Similarity 64.3%; Pred. No. 84;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 VLLGKLLWKIDNPG 17
 DB 15 VLLGSLWPLAAPG 28
 Search completed: November 10, 2004, 14:50:08
 Job time : 25.9774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 8.05283 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-50

Perfect score: 111
Sequence: 1 GAMEREWAMFLRAASSRRGGV 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	41.4	250	1	US-08-378-761A-71
2	46	41.4	250	1	US-08-485-286-71
3	46	41.4	251	4	US-09-538-873-3
4	44	39.6	151	4	US-09-489-039A-9109
5	43	38.7	225	4	US-09-538-092-962
6	43	38.7	636	4	US-09-252-991A-22829
7	42.5	38.3	234	4	US-09-134-000C-5253
8	42	37.8	200	4	US-09-538-092-249
9	42	37.8	671	3	US-08-858-207A-301
10	42	37.8	671	4	US-09-583-110-3642
11	42	37.8	1464	4	US-10-038-224-2
12	42	37.8	1614	3	US-09-052-469-2
13	42	37.8	1614	4	US-08-422-582-2
14	42	37.8	1614	4	US-09-052-282-2
15	42	37.8	4302	3	US-08-658-136-5
16	42	37.8	4302	3	US-09-052-469-8
17	42	37.8	4302	4	US-08-422-582-8
18	42	37.8	4302	4	US-09-052-262-8
19	42	37.8	4303	2	US-08-460-751-2
20	42	37.8	4303	4	US-09-479-457A-2
21	42	37.8	4339	3	US-09-052-469-6
22	42	37.8	4339	4	US-08-422-582-6
23	42	37.8	4339	4	US-09-052-262-6
24	41	36.9	153	4	US-09-252-991A-32617
25	41	36.9	395	4	US-09-252-991A-31694
26	40	36.0	18	4	US-09-971-806-2
27	40	36.0	138	4	US-09-252-991A-26931

28	40	36.0	185	4	US-09-252-991A-27468	Sequence 27468, A
29	40	36.0	292	4	US-09-252-991A-28750	Sequence 28750, A
30	40	36.0	323	4	US-09-252-991A-21180	Sequence 21180, A
31	40	36.0	418	4	US-09-252-991A-32633	Sequence 32633, A
32	40	36.0	487	4	US-09-248-796A-19788	Sequence 19788, A
33	40	36.0	921	1	US-08-188-582-2	Sequence 2, Appli
34	40	36.0	921	1	US-08-646-715-2	Sequence 2, Appli
35	39.5	35.6	2190	4	US-09-252-991A-25754	Sequence 25754, A
36	39	35.1	303	3	US-08-660-645A-3	Sequence 3, Appli
37	39	35.1	303	3	US-09-298-718-3	Sequence 3, Appli
38	39	35.1	303	3	US-09-546-969-3	Sequence 3, Appli
39	39	35.1	303	3	US-08-980-832-3	Sequence 3, Appli
40	39	35.1	303	3	US-09-547-267-3	Sequence 3, Appli
41	39	35.1	303	4	US-09-920-923B-3	Sequence 3, Appli
42	39	35.1	322	4	US-09-252-991A-31782	Sequence 31782, A
43	39	35.1	345	4	US-09-489-039A-8306	Sequence 8306, Ap
44	39	35.1	620	3	US-08-982-785A-10	Sequence 10, Appl
45	39	35.1	620	4	US-09-629-498-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-378-761A-71
; Sequence 71, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,761A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-378-761A-71

Query Match 41.4%; Score 46; DB 1; Length 250;
Best Local Similarity 36.4%; Pred. No. 7.4;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 1 GAMEREWAMFLRAASSRRGGV 22
||| : : : : :
Db 11 GATSQSYKQFIEALRELRGGL 32

```
RESULT 2
US-08-485-286-71
; Sequence 71, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 08/378761
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-286-71

Query Match 41.4%; Score 46; DB 1; Length 250;
Best Local Similarity 36.4%; Pred. No. 7.4;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMREWAMFLRAASSRRGGV 22
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 3
US-09-538-873-3
; Sequence 3, Application US/09538873
; Patent No. 6566500
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETTIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
; TITLE OF INVENTION: SYNDROME (VLS)
; FILE REFERENCE: UTSD:603
; CURRENT APPLICATION NUMBER: US/09/538,873
; CURRENT FILING DATE: 2000-03-30
; EARLIER APPLICATION NUMBER: 60/126,826
```

```
; EARLIER FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Abrus precatorius
; US-09-538-873-3

Query Match 41.4%; Score 46; DB 4; Length 251;
Best Local Similarity 36.4%; Pred. No. 7.5;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMREWAMFLRAASSRRGGV 22
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 4
US-09-489-039A-9109
; Sequence 9109, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9109
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-9109

Query Match 39.6%; Score 44; DB 4; Length 151;
Best Local Similarity 50.0%; Pred. No. 9.1;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 REWAMFLRAASSRRG 20
DB 34 RTWVLEFMAGASRQQ 49

RESULT 5
US-09-538-092-862
; Sequence 862, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqformat Version 0.9
; SEQ ID NO 862
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P07029
; US-09-538-092-862
```

```
Query Match      38.7%; Score 43; DB 4; Length 225;
Best Local Similarity 38.9%; Pred. No. 21;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 AMERWAMFLRAASSRIR 19
DB 90 AQBREWKFIRALAGTLQ 107

RESULT 6
US-09-252-991A-22829
; Sequence 22829, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22829
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22829

Query Match      38.7%; Score 43; DB 4; Length 636;
Best Local Similarity 50.0%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 REWAMFLRAASSRRING 20
DB 557 REWAVFLRRVFQAHG 572

RESULT 7
US-09-134-000C-5253
; Sequence 5253, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5253
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5253

Query Match      38.3%; Score 42.5; DB 4; Length 234;
Best Local Similarity 44.0%; Pred. No. 27;
Matches 11; Conservative 4; Mismatches 3; Indels 7; Gaps 2;

QY 3 MEREWAMEL----RAASS---RIRG 20
DB 11 MEREWELFLAPYEQAVSELKVLRG 35

RESULT 8
US-09-538-092-249
```

```
; Sequence 249, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqformatter Version 0.9
; SEQ ID NO 249
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YER102W
US-09-538-092-249

Query Match      37.8%; Score 42; DB 4; Length 200;
Best Local Similarity 52.6%; Pred. No. 27;
Matches 10; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 4 EREWAMFLRAASSRRINGV 22
DB 140 ERKWA--ARAASAKTIESSV 156

RESULT 9
US-08-858-207A-301
; Sequence 301, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R.
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
```



```

US-08-422-582-2
; Sequence 2, Application US/08422582
; Patent No. 6485960
; GENERAL INFORMATION:
; APPLICANT: Harris et al.
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE 1 GENE
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 75 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.50 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,582
; FILING DATE: 14-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9411900.5
; FILING DATE: 14-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB PCT/GB94/02822
; FILING DATE: 23-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326470.3
; FILING DATE: 24-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 3265/53313 (MRC-006xx)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1614 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-422-582-2

Query Match 37.8%; Score 42; DB 4; Length 1614;
Best Local Similarity 41.2%; Pred. No. 2.8e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GAMERWAMFLRAASSR 17
Db 1188 GAADQWTFVGRPRR 1204

RESULT 14
US-09-052-262-2
; Sequence 2, Application US/09052262
; Patent No. 6656681
; GENERAL INFORMATION:
; APPLICANT: Harris et al.
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE 1 GENE
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: One Financial Center
; CITY: Boston
; STATE: MA

```

```

; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.50 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/052,262
; FILING DATE: Concurrently herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/422,582
; FILING DATE: 14-April-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9507766.5
; FILING DATE: 13-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9411900.5
; FILING DATE: 14-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB PCT/GB94/02822
; FILING DATE: 23-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326470.3
; FILING DATE: 24-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 3265/74118
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1614 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-052-262-2

Query Match 37.8%; Score 42; DB 4; Length 1614;
Best Local Similarity 41.2%; Pred. No. 2.8e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GAMERWAMFLRAASSR 17
Db 1188 GAADQWTFVGRPRR 1204

RESULT 15
US-08-658-136-5
; Sequence 5, Application US/08658136
; Patent No. 6071717
; GENERAL INFORMATION:
; APPLICANT: KLINGER, KATHERINE W
; APPLICANT: LANDES, GREGORY M
; APPLICANT: BUEN, TIMOTHY C
; APPLICANT: CONNORS, TIMOTHY D
; APPLICANT: DACKOWSKI, WILLIAM
; APPLICANT: GERMINO, GREGORY
; APPLICANT: QIAN, FENG
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: ONE MOUNTAIN ROAD
; CITY: FRAMINGHAM
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/658,136
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LASSEN, ELIZABETH
; REGISTRATION NUMBER: 31,845
; REFERENCE/DOCKET NUMBER: GEN4-17.8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 508-872-8400
; TELEFAX: 508-872-5415
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-658-136-5

Query Match      37.8%; Score 42; DB 3; Length 4302;
Best Local Similarity 41.2%; Pred. No. 8.5e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      1 GAMEREMAMFLRAASSR 17
Db      3956 GAADROWTRFVRGPRR 3972

```

Search completed: November 10, 2004, 14:55:37
Job time : 9.10283 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 25.4453 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-50

Perfect score: 111

Sequence: 1 GAMERWAMFLRAASSRIRGGV 22

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 35322586 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	111	100.0	22	14 US-10-092-750-50	Sequence 50, Appl
2	98	88.3	82	13 US-10-001-835-218	Sequence 218, App
3	98	88.3	82	15 US-10-001-885-107	Sequence 107, App
4	48.5	43.7	176	15 US-10-389-566-958	Sequence 958, App
5	47.5	42.8	204	17 US-10-425-115-224673	Sequence 224673,
6	47	42.3	92	15 US-10-424-599-270202	Sequence 270202,
7	46	41.4	189	15 US-10-424-599-209333	Sequence 209333,
8	46	41.4	251	14 US-10-282-935-3	Sequence 3, Appli
9	46	41.4	251	15 US-10-440-796-3	Sequence 3, Appli
10	46	41.4	434	16 US-10-437-963-190381	Sequence 190381,
11	45.5	41.0	328	14 US-10-152-886-37	Sequence 37, Appl
12	45	40.5	170	17 US-10-425-115-331184	Sequence 331184,
13	44	39.6	43	15 US-10-424-599-263656	Sequence 263656,

14	44	39.6	192	17	US-10-425-115-325061	Sequence 325061,
15	44	39.6	287	16	US-10-767-701-43701	Sequence 43701, A
16	44	39.6	290	15	US-10-282-122A-49678	Sequence 49678, A
17	44	39.6	359	16	US-10-437-963-108434	Sequence 108434,
18	44	39.6	464	16	US-10-437-963-151611	Sequence 151611,
19	44	39.6	468	17	US-10-425-115-323337	Sequence 323337,
20	44	39.6	508	16	US-10-437-963-154485	Sequence 154485,
21	44	39.6	602	14	US-10-104-047-3324	Sequence 3324, Ap
22	44	39.6	1100	16	US-10-437-963-196344	Sequence 196344,
23	43	38.7	128	17	US-10-425-115-271121	Sequence 271121,
24	43	38.7	211	17	US-10-425-115-340893	Sequence 340893,
25	43	38.7	278	16	US-10-767-701-32056	Sequence 32056, A
26	43	38.7	357	15	US-10-425-114-60544	Sequence 60544, A
27	43	38.7	374	15	US-10-282-122A-70161	Sequence 70161, A
28	43	38.7	555	17	US-10-425-115-227829	Sequence 227829,
29	43	38.7	582	15	US-10-425-114-54127	Sequence 54127, A
30	43	38.7	1803	16	US-10-343-509-3	Sequence 3, Appli
31	43	38.7	1805	16	US-10-343-509-3	Sequence 1, Appli
32	42.5	38.3	125	17	US-10-425-115-190434	Sequence 190434,
33	42.5	38.3	224	15	US-10-282-122A-42466	Sequence 42466, A
34	42.5	38.3	234	9	US-09-815-242-10620	Sequence 10620, Ap
35	42.5	38.3	501	14	US-10-094-749-2824	Sequence 2824, Ap
36	42.5	38.3	559	16	US-10-437-963-128120	Sequence 128120,
37	42	37.8	72	15	US-10-424-599-179626	Sequence 179626,
38	42	37.8	81	15	US-10-424-599-243709	Sequence 243709,
39	42	37.8	132	16	US-10-437-963-147013	Sequence 147013,
40	42	37.8	142	17	US-10-425-115-308921	Sequence 308921,
41	42	37.8	149	15	US-10-424-599-147227	Sequence 147227,
42	42	37.8	180	16	US-10-437-963-203121	Sequence 203121,
43	42	37.8	200	16	US-10-451-467A-28	Sequence 28, Appl
44	42	37.8	213	16	US-10-437-963-181988	Sequence 181988,
45	42	37.8	534	14	US-10-156-761-12074	Sequence 12074, A

ALIGNMENTS

RESULT 1
US-10-092-750-50
; Sequence 50, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/05002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-50

Query Match 100.0%; Score 111; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GAMERWAMFLRAASSRIRGGV 22
Db 1 GAMERWAMFLRAASSRIRGGV 22

RESULT 2
US-10-001-835-218
; Sequence 218, Application US/10001835
; Publication No. US20020160387A1
; GENERAL INFORMATION:

```

; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Caferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
; CURRENT APPLICATION NUMBER: US/10/001,835
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249,997
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 218
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-835-218

Query Match      88.3%; Score 98; DB 13; Length 82;
Best Local Similarity 90.9%; Pred. No. 6.2e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 36 GDMEREWAMFLRAASSGIRGGV 57

RESULT 3
US-10-001-885-107
; Sequence 107, Application US/10001885
; Publication No. US20040058319A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Caferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
; CURRENT APPLICATION NUMBER: US/10/001,885
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,061
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/253,257
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 107
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-885-107

Query Match      88.3%; Score 98; DB 15; Length 82;
Best Local Similarity 90.9%; Pred. No. 6.2e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 36 GDMEREWAMFLRAASSGIRGGV 57

RESULT 4
US-10-389-566-958
; Sequence 958, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants

```

```

; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 958
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (43)..(44)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65)..(65)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (165)..(165)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-389-566-958

Query Match      43.7%; Score 48.5; DB 15; Length 176;
Best Local Similarity 57.9%; Pred. No. 12;
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 1 GAMEREWAMFLRAASSRI 18
DB 141 GARWREWIAMLLRTSTSRV 159

RESULT 5
US-10-425-115-224673
; Sequence 224673, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 224673
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_13648C.1.pep
US-10-425-115-224673

Query Match      42.8%; Score 47.5; DB 17; Length 204;
Best Local Similarity 48.0%; Pred. No. 20;
Matches 12; Conservative 2; Mismatches 6; Indels 5; Gaps 1;

QY 2 AMER-----EWAMFLRAASSRIRGG 21
DB 112 AAEREVAVWWSLLRSVAVGRHGG 136

```

```

; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF
; TITLE OF INVENTION: PROTEINACEOUS COMPOUNDS
; FILE REFERENCE: UTSD:884US
; CURRENT APPLICATION NUMBER: US/10/282,935
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 09/538,873
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/126,826
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Abrus precatorius
US-10-282-935-3

Query Match          41.4%; Score 46; DB 14; Length 251;
Best Local Similarity 36.4%; Pred.No. 42;
Matches      8; Conservative    5; Mismatches   9; Indels    0; Gaps    0

QY      1 GAMEREWAMFLRAASSRIRGGV 22
       |||...|||...|||:
Db      11 GATSQYKQFIALRRLRGGL 32

RESULT 9
US-10-440-796-3
; Sequence 3, Application US/10440796
; Publication No. US20040009148A1
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
; TITLE OF INVENTION: SYNDROME (VLS)
; FILE REFERENCE: UTSD:603
; CURRENT APPLICATION NUMBER: US/10/440,796
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US/09/538,873
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/126,826
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Abrus precatorius
US-10-440-796-3

Query Match          41.4%; Score 46; DB 15; Length 251;
Best Local Similarity 36.4%; Pred.No. 42;
Matches      8; Conservative    5; Mismatches   9; Indels    0; Gaps    0

QY      1 GAMEREWAMFLRAASSRIRGGV 22
       |||...|||...|||:
Db      11 GATSQYKQFIALRRLRGGL 32

RESULT 10
US-10-437-963-190381
; Sequence 190381, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
```

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 190381
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86801C.1.pep
US-10-437-963-190381

Query Match 41.4%; Score 46; DB 16; Length 434;
Best Local Similarity 50.0%; Pred. No. 74;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 MEREWAMFLRASSRI 20
Db 390 VERMWSFIRAGGSRIPG 407

RESULT 11
US-10-152-886-37
; Sequence 37, Application US/10152886
; Publication No. US20030064491A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: Farnet, Chris
; APPLICANT: Staffa, Alfredo
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: GENES AND PROTEINS INVOLVED IN THE BIOSYNTHESIS OF ENEDIYNE RING
; FILE REFERENCE: 3011-305
; CURRENT APPLICATION NUMBER: US/10/152,886
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Streptomyces carzinostaticus neocarzinostaticus
US-10-152-886-37

Query Match 41.0%; Score 45.5; DB 14; Length 328;
Best Local Similarity 32.4%; Pred. No. 67;
Matches 12; Conservative 5; Mismatches 5; Indels 15; Gaps 1;

QY 1 GAMEREWAMFLR-----AASRI 22
Db 229 GATEQELALFARQAGEHRAWALAQGAFAAEARVRAGL 265

RESULT 12
US-10-425-115-331184
; Sequence 331184, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 331184
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_65144C.1.pep
US-10-425-115-331184

Query Match 40.5%; Score 45; DB 17; Length 170;
Best Local Similarity 40.7%; Pred. No. 41;
Matches 11; Conservative 3; Mismatches 7; Indels 6; Gaps 1;

QY 1 GAMEREWAMFLRA-----ASSRI 21
Db 114 GRRKRWERTMAVAVADACASSRLRG 140

RESULT 13
US-10-424-599-263656
; Sequence 263656, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285884
; SEQ ID NO 263656
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80100C.1.pep
US-10-424-599-263656

Query Match 39.6%; Score 44; DB 15; Length 43;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 9; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

QY 5 REWAMFLRASSRI--RGG 21
Db 25 REWSLMFKASSTLVDRGG 43

RESULT 14
US-10-425-115-325061
; Sequence 325061, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 325061
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_59526C.1.pep
US-10-425-115-325061

```
Query Match      39.6%; Score 44; DB 17; Length 192;
Best Local Similarity 50.0%; Pred. No. 68;
Matches      8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      5 REMAMFLRAASSRIQG 20
DB      135 RRWNMLEAENSLLRG 150

RESULT 15
US-10-767-701-43701
; Sequence 43701, Application US/10767701
; Publication No: US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 43701
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C13563_1.pep
US-10-767-701-43701

Query Match      39.6%; Score 44; DB 16; Length 287;
Best Local Similarity 61.5%; Pred. No. 1e+02;
Matches      8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 GAMEREWAMFLRA 13
DB      240 GTMEQRWQAQFERA 252

Search completed: November 11, 2004, 02:43:06
Job time : 26.4953 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 5.39623 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-50

Perfect score: 111
Sequence: 1 GAMERWAFLLRAASSRRGGV 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	41.4	249	2 C84185	hypothetical prote
2	46	41.4	251	2 C39761	abrin (clone 7.2)
3	46	41.4	528	1 T2LSA	abrin-a precursor
4	45	40.5	231	2 G3823	hypothetical prote
5	45	40.5	266	2 A2701	aldolase Atul014
6	45	40.5	266	2 G37483	hypothetical prote
7	44	39.6	488	1 E69473	probable DNA-direc
8	44	39.6	1166	2 T15628	hypothetical prote
9	43	38.7	154	2 E82112	conserved hypotet
10	43	38.7	225	2 A24016	heterogeneous ribo
11	43	38.7	399	2 B93069	GTP-binding protei
12	43	38.7	657	2 A3170	conserved hypotet
13	43	38.7	833	2 H72205	maltose ABC transp
14	42.5	38.3	458	2 H75603	conserved hypotet
15	42.5	38.3	2139	2 A35672	crumbs protein - f
16	42	37.8	200	2 S45591	ribosomal protein
17	42	37.8	302	2 A70855	probable transfera
18	42	37.8	521	2 T01923	hypothetical prote
19	42	37.8	541	2 A24933	hypothetical prote
20	42	37.8	811	2 S57149	probable membrane
21	42	37.8	1017	2 S64782	probable membrane
22	42	37.8	1040	2 T08190	hypothetical prote
23	42	37.8	4302	2 A38971	polycystic kidney
24	41	36.9	149	2 A80812	conserved hypotet
25	41	36.9	330	2 T20057	hypothetical prote
26	41	36.9	336	2 T22285	hypothetical prote
27	41	36.9	458	2 B82216	para-aminobenzoate
28	41	36.9	605	2 S25151	invertase, cell-wa
29	41	36.9	654	2 B55579	biotin carboxyl ca

30 41 36.9 1110 1 A70652 probable serine/th
31 41 36.9 1706 2 B75633 probable RNA helic
32 40.5 36.5 377 2 T31604 hypothetical prote
33 40 36.0 229 2 E72730 hypothetical prote
34 40 36.0 230 2 S29158 hypothetical prote
35 40 36.0 243 2 E87413 hemolysin A [impor
36 40 36.0 285 2 E83257 tRNA-pseudouridine
37 40 36.0 324 2 A69290 probable glycosyl
38 40 36.0 346 2 E95870 probable dTPPGluco
39 40 36.0 371 2 A39234 opsin - bluebottle
40 40 36.0 434 2 A96609 probable polygalac
41 40 36.0 447 2 T45653 hypersensitivity r
42 40 36.0 485 2 T35663 probable transmem
43 40 36.0 518 2 AC0675 hypothetical prote
44 40 36.0 527 2 S32430 abrin-b precursor
45 40 36.0 528 2 S32431 abrin-d precursor

ALIGNMENTS

RESULT 1

C84185
hypothetical protein Vng0250c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Dec-2002
C;Accession: C84185
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jab
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.;
A;Title: Genome sequence of Halobacterium species NRC-1
A;Reference number: A84160; MUID:20504483; PMID:11018950
A;Accession: C84185
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <STO>
A;Cross-references: GB:AE004437; NID:gl0579889; PIDN:AGI18847.1; GSPDB:GNO0138
C;Genetics:
A;Gene: VNG0250C
C;Superfamily: uncharacterized conserved protein

Query Match 41.4%; Score 46; DB 2; Length 249;
Best Local Similarity 63.2%; Pred. No. 5.6;
Matches 12; Conservative 2; Mismatches 3; Indels 2; Gaps 1;
QY 4 BREWAFLLRAASSRRGGV 22
: ||| ||| ||| ||| :
Db 229 DREAA--LRAAEERINGGM 245

RESULT 2

C39761
abrin (clone 7.2) precursor - Indian licorice (fragment)
N;Contains: rRNA N-glycosidase (EC 3.2.2.22)
C;Species: Abrus precatorius (Indian licorice)
C;Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: C39761; S14471
R;Svensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266, 6848-6852, 1991
A;Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A;Reference number: A39761; MUID:91201329; PMID:2016300
A;Accession: C39761
A;Molecule type: DNA
A;Residues: 1-251 <EVE>
A;Cross-references: UNIPROT:Q38760
R;Svensen, G.; Mathiesen, A.; Sundan, A.
submitted to the EMBL Data Library, October 1990
A;Description: Direct molecular cloning of two distinct abrin A-chains.
A;Reference number: S14471
A;Accession: S14471
A;Molecule type: DNA

A;Residues: 'M', 1-251 <EV2>
A;Cross-references: EMBL:X54872; NID:gl6088; PIDN:CAA38654.1; PID:gl6089
C;Superfamily: ricin; rRNA N-glycosidase homology
C;Keywords: duplication; abrin (clone 7.2) chain A #status predicted <ACH>
F;1-251/Product: abrin (clone 7.2) chain A #status predicted <ACH>
F;7-246/Domain: rRNA N-glycosidase homology <RNG>
F;74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F;164,167/Active site: Glu, Arg #status predicted

Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 5.6;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMERWAMFLRAASRRIGGV 22
DB 11 GATSOYKQFIALRRLRGL 32

RESULT 3
TZLSA
abrin-a precursor - Indian licorice (fragment)
N;Contains: rRNA N-glycosidase (EC 3.2.2.22)
C;Species: Abrus precatorius (Indian licorice)
C;Date: 31-Dec-1993 #sequence revision 01-Aug-1997 #text change 09-Jul-2004
C;Accession: S32429; J0202; A39761; J01398; S14472; S24133; S74110; S74111
R;Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J. Mol. Biol. 229, 263-267, 1993
A;Title: Primary structure of three distinct isoabrin determined by cDNA sequencing. C
A;Reference number: S32429; MUID:93132798; PMID:8421313
A;Accession: S32429
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 'E', 2-528 <HUN>
A;Cross-references: UNIPROT:P11140; GB:M98344; NID:gl66294; PIDN:AAA32624.1; PID:gl66295
A;Note: the coding region for the sequence shown is preceded by an ATG codon
A;Note: residues 1-8 were derived from the synthesized primer
R;Funatsu, G.; Taguchi, Y.; Kamemoto, M.; Yanaka, M.
Agric. Biol. Chem. 52, 1095-1097, 1988
A;Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein fr
A;Reference number: J0202
A;Accession: J0202
A;Molecule type: protein
A;Status: nucleic acid sequence not shown
A;Residues: 1-201,203-251 <FUN>
A;Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
R;Evensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266, 6848-6852, 1991
A;Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A;Reference number: A39761; MUID:91201329; PMID:2016300
A;Accession: A39761
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 'E', 2-251 <EV5>
A;Cross-references: GB:X54872
A;Note: residues 1-8 were derived from the synthesized primer
R;Kimura, M.; Sumizawa, T.; Funatsu, G.
Biosci. Biotechnol. Biochem. 57, 166-169, 1993
A;Title: The complete amino acid sequences of the B-chains of abrin-a and abrin-b, toxic
A;Reference number: J01398; MUID:93169023; PMID:7763422
A;Contents: seeds
A;Accession: J01398
A;Molecule type: protein
A;Residues: 261-347, 'T', 349-351, 'A', 353-357, 'L', 359-528 <KIM>
A;Experimental source: seed
R;Evensen, G.; Mathiesen, A.; Sundan, A.
submitted to the EMBL data library, October 1990
A;Description: Direct molecular cloning of two distinct abrin A-chains.
A;Reference number: S14471
A;Accession: S14472
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 'ME', 2-251 <BV2>
A;Cross-references: EMBL:X54873; NID:gl6090; PIDN:CAA38655.1; PID:gl6091
R;Chen, Y.L.; Chow, L.P.; Tsugita, A.; Lin, J.Y.

FERS Lett. 309, 115-118, 1992
A;Title: The complete primary structure of abrin-a B chain.
A;Reference number: S24133; MUID:92371656; PMID:1505674
A;Accession: S24133
A;Molecule type: protein
A;Residues: 262-297, 'Y', 299-426, 'L', 428-466, 'P', 468-482, 'L', 484-528 <CHE>
R;Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.
Eur. J. Biochem. 240, 564-569, 1996
A;Title: Probing the domain structure of abrin-a by tryptic digestion.
A;Reference number: S74110; MUID:97008945; PMID:8856055
A;Accession: S74110
A;Molecule type: protein
A;Residues: 89-108;154-172 <LIN>
A;Experimental source: seed
A;Accession: S74111
A;Molecule type: protein
A;Residues: 262-276, 'X', 278-280;329-348;369-388;399-418 <LIW>
A;Experimental source: seed
C;Comment: Abrin-a is more toxic than ricin. The toxin consists of an A chain, which in
taining receptors on the cell surface. The A and B chains are linked by a single disulf
C;Superfamily: ricin; rRNA N-glycosidase homology
C;Keywords: duplication; abrin-a chain A #status experimental <ACH>
F;1-251/Product: abrin-a chain A #status experimental <ACH>
F;7-246/Domain: rRNA N-glycosidase homology <RNG>
F;261-528/Product: abrin-a chain B #status experimental <BOH>
F;283-325;326-366;369-407;414-449;453-492;495-528/Region: 40-residue repeats
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F;164,167/Active site: Glu, Arg #status predicted
F;247-269;286-305;329-346;417-430;456-473/Disulfide bonds: #status predicted
F;288,312/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted
F;361,401/Binding site: carboxylate (Asn) (covalent) #status experimental
F;500,521/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 41.4%; Score 46; DB 1; Length 528;
Best Local Similarity 36.4%; Pred. No. 12;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMERWAMFLRAASRRIGGV 22
DB 11 GATSOYKQFIALRRLRGL 32

RESULT 4
G83823
Hypothetical protein BH1391 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: G83823
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A3650; MUID:20512582; PMID:11058132
A;Accession: G83823
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-231 <STO>
A;Cross-references: UNIPROT:Q9KD28; GB:AP001511; GB:BA000004; NID:gl0173727; PIDN:BA05
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1391

Query Match 40.5%; Score 45; DB 2; Length 231;
Best Local Similarity 47.1%; Pred. No. 7.6;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAMERWAMFLRAASRR 17
DB 83 GAFKEWGWFLAGISQQ 99

RESULT 5
AF2701

aidolase Atu1014 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C;Species: Agrobacterium tumefaciens
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C;Accession: AF2701
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A;Reference number: AB2577; MUID:21608550; PMID:11743193
 A;Accession: AF2701
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-266 <KUR>
 A;Cross-references: UNIPROT:Q8UGM2; GB:AE008688; PIDN:AAL42028.1; PID:q17739404; GSPDB:G
 A;Experimental source: strain C58 (Dupont)
 C;Genetics:
 A;Gene: Atu1014
 A;Map position: circular chromosome
 C;Superfamily: 2,4-dihydroxyhept-2-ene-1,7

Query Match 40.5%; Score 45; DB 2; Length 266;
 Best Local Similarity 52.4%; Pred. No. 8.7;
 Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAMERWAMFLRAASSRIRGG 21
 |||||
 Db 236 GADVTEFANTLRALSARYKGG 256

RESULT 6

G97483
 hypothetical protein AGR_C1872 [imported] - Agrobacterium tumefaciens (strain C58, Cere
 C;Species: Agrobacterium tumefaciens
 C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C;Accession: G97483
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A;Reference number: A97359; MUID:21608551; PMID:11743194
 A;Accession: G97483
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-266 <KUR>
 A;Cross-references: UNIPROT:Q8UGM2; GB:AE007869; PIDN:AAK86824.1; PID:g15156034; GSPDB:G
 C;Genetics:
 A;Gene: AGR_C1872
 A;Map position: circular chromosome
 C;Superfamily: 2,4-dihydroxyhept-2-ene-1,7

Query Match 40.5%; Score 45; DB 2; Length 266;
 Best Local Similarity 52.4%; Pred. No. 8.7;
 Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAMERWAMFLRAASSRIRGG 21
 |||||
 Db 236 GADVTEFANTLRALSARYKGG 256

RESULT 7

E69473
 probable DNA-directed DNA polymerase (EC 2.7.7.7) delta small chain - Archaeoglobus fulg
 N;Alternate names: conserved hypothetical protein AF1930
 C;Species: Archaeoglobus fulgidus
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C;Accession: E69473
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997

A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
 A;Reference number: A69250; MUID:198049343; PMID:9389475
 A;Accession: E69473
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-488 <KLE>
 A;Cross-references: UNIPROT:Q28484; GB:AE000979; GB:AE000782; NID:g2689302; PIDN:AA8894
 C;Superfamily: DNA polymerase II small subunit, archaeal type; phosphoesterase core hom
 C;Keywords: metalloprotein; nucleotidyltransferase
 F;228-326/Domain: phosphoesterase core homology >PEC>

Query Match 39.6%; Score 44; DB 1; Length 488;
 Best Local Similarity 38.9%; Pred. No. 23;
 Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 MEREWAMFLRAASSRIRGG 20
 :|||
 Db 243 LEKEMEMFVRWLKGEVGG 260

RESULT 8

T15628
 hypothetical protein C25H3.9 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C;Accession: T15628
 R;Johnson, D.
 A;Submitted to the EMBL Data Library, June 1995
 A;Description: The sequence of C. elegans cosmid C25H3.
 A;Reference number: Z18379
 A;Accession: T15628
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1166 <JOU>
 A;Cross-references: EMBL:U29535; NID:g868251; PID:g868258; PIDN:AAA68787.1; CESP:C25H3.
 A;Experimental source: strain Bristol N2
 C;Genetics:
 A;Gene: CESP:C25H3.9
 A;Introns: 45/3; 82/3; 129/2; 180/2; 202/1; 215/3; 252/1; 374/3; 544/2; 1028/3; 1108/1

Query Match 39.6%; Score 44; DB 2; Length 1166;
 Best Local Similarity 56.2%; Pred. No. 56;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 AMERWAMFLRAASSR 17
 |||||
 Db 518 AVNREWARPERARQHR 533

RESULT 9

E82112
 conserved hypothetical protein VC2146 [imported] - Vibrio cholerae (strain N16961 serog
 C;Species: Vibrio cholerae
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C;Accession: E82112
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers,
 I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A;Reference number: AB2035; MUID:20406833; PMID:10952301

A;Accession: E82112
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-154 <HEI>
 A;Cross-references: UNIPROT:Q9KQ58; GB:AE004287; GB:AE003852; NID:g9656689; PIDN:AAF952
 A;Experimental source: serogroup O1, strain N16961; biotype El Tor
 C;Genetics:
 A;Gene: VC2146
 A;Map position: 1

Query Match 38.7%; Score 43; DB 2; Length 154;
 Best Local Similarity 44.4%; Pred. No. 11;
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 MEREWAMFLRAASSRIRG 20
 : ||||| :
 Db 27 LNRWVWFVMAVSREQ 44

RESULT 10
 A24016
 heterogeneous ribonuclear particle protein C - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 09-Jul-2004
 C:Accession: A24016
 R: Lahiri, D.K.; Thomas, J.O.
 Nucleic Acids Res. 14, 4077-4094, 1986
 A:Title: A cDNA clone of the hRNP C proteins and its homology with the single-stranded
 A:Reference number: A24016; MUID:86232544; PMID:3754960
 A:Accession: A24016
 A:Molecule type: mRNA
 A:Residues: 1-225 <LAH>
 A:Cross-references: UNIPROT:P07029; GB:X03910; NID:G36103; PIDN:CAA27544.1; PID:G929663
 C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
 C:Keywords: DNA binding

Query Match 38.7%; Score 43; DB 2; Length 225;
 Best Local Similarity 38.9%; Pred. No. 16;
 Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 AMEREWAMFLRAASSRIR 19
 : ||||| :
 Db 90 AQREWKQFIRALAGTLQ 107

RESULT 11
 B69069
 GTP-binding protein - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: B69069
 R: Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qiu, D.; Spadafora, R.; Vicaito, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: B69069
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-399 <MTH>
 A:Cross-references: UNIPROT:Q27559; GB:AE000912; GB:AE000666; NID:G2622632; PIDN:AAH8599
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1515
 A:Start codon: GTG
 C:Superfamily: Methanococcus jannaschii GTP-binding protein
 C:Keywords: GTP binding; nucleotide binding; P-loop
 F:11-18/Region: nucleotide-binding motif A (P-loop)
 F:226-229/Region: GTP-binding NKXD motif

Query Match 38.7%; Score 43; DB 2; Length 399;
 Best Local Similarity 50.0%; Pred. No. 28;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASS 16
 : ||||| :
 Db 150 GILERNWERLLKKAAS 165

RESULT 12
 AG3170
 conserved hypothetical protein Atu5090 [imported] - Agrobacterium tumefaciens (strain CS

C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C:Accession: AG3170
 R: Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AG3170
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-657 <KUR>
 A:Cross-references: UNIPROT:Q8UKL1; GB:AE008697; PIDN:AAL45781.1; PID:G17743516; GSPDB:
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu5090
 A:Genome: plasmid
 C:Superfamily: bcsA protein

Query Match 38.7%; Score 43; DB 2; Length 657;
 Best Local Similarity 40.0%; Pred. No. 46;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRG 20
 : ||||| :
 Db 561 GALKRVWMTLQETARLRG 580

RESULT 13
 H72205
 maltose ABC transporter, permease protein - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C:Accession: H72205
 R: Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hicke
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: H72205
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-833 <ARN>
 A:Cross-references: UNIPROT:Q3X2P5; GB:AE001820; GB:AE000512; NID:G4982411; PIDN:AAD368
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM1836

Query Match 38.7%; Score 43; DB 2; Length 833;
 Best Local Similarity 57.1%; Pred. No. 59;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAA 14
 : ||||| :
 Db 791 GAYETEWGLFTAAA 804

RESULT 14
 H75603
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: H75603
 R: White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; M
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 29.3057 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-50
Perfect score: 111
Sequence: 1 GAMERWAMFLRAASSRIRGGV 22

Scoring table: BLCOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	46	41.4	249	2	Q9HSG0	Q9hsg0 halobacteri	
2	46	41.4	251	2	Q7DM12	Q7dm12 abrus preca	
3	46	41.4	251	2	Q96235	Q96235 abrus preca	
4	46	41.4	251	2	Q96236	Q96236 abrus preca	
5	46	41.4	251	2	Q96237	Q96237 abrus preca	
6	46	41.4	252	2	Q38760	Q38760 abrus preca	
7	46	41.4	252	2	Q38761	Q38761 abrus preca	
8	46	41.4	319	2	Q88GL5	Q88GL5 pseudomonas	
9	46	41.4	382	2	Q8VVD3	Q8vyd3 arabidopsis	
10	46	41.4	528	1	ABRA_ABRPR	Pili140 abrus preca	
11	45.5	41.0	328	2	Q83U65	Q83u65 streptomyces	
12	45	40.5	231	2	Q9KD28	Q9kd28 bacillus halodurans	
13	45	40.5	266	2	Q7D052	Q7d052 agrobacteri	
14	45	40.5	266	2	Q8UGM2	Q8ugm2 agrobacteri	
15	45	40.5	1423	2	Q8EVP5	Q8evp5 yersinia pseudotuberculosis	
16	44	39.6	215	2	Q8ZFE6	Q8zfe6 oryza sativa	
17	44	39.6	215	2	BAD09297	BAD09297 oryza sativa	
18	44	39.6	219	2	Q6YTE0	Q6yte0 oryza sativa	
19	44	39.6	219	2	BAD17669	BAD17669 oryza sativa	
20	44	39.6	219	2	BAD17797	BAD17797 oryza sativa	
21	44	39.6	224	2	Q84NH7	Q84nh7 manihot esculenta	
22	44	39.6	224	2	Q84NH6	Q84nh6 manihot esculenta	
23	44	39.6	230	2	Q9DKW4	Q9dkw4 cucurbit ficus	
24	44	39.6	258	2	Q7NNR6	Q7nnr6 gloeobacterium	
25	44	39.6	288	2	Q7X2L4	Q7x2l4 sphingomonas	
26	44	39.6	337	2	Q8H9F3	Q8h9f3 oryza sativa	
27	44	39.6	413	2	Q87HU6	Q87hu6 vibrio parvulus	
28	44	39.6	416	2	Q7YWB4	Q7ywb4 drosophila	
29	44	39.6	416	2	Q9V7J4	Q9v7j4 drosophila	
30	44	39.6	437	2	Q8MLC9	Q8mlc9 drosophila	
31	44	39.6	448	2	Q8IHB4	Q8ihb4 drosophila	

32	44	39.6	464	2	Q7XCS3	Q7xcs3 oryza sativ
33	44	39.6	464	2	Q94HY3	Q94hy3 oryza sativ
34	44	39.6	481	2	Q6ZGA1	Q6zga1 oryza sativ
35	44	39.6	481	2	BAD07736	BAD07736 oryza sat
36	44	39.6	488	1	DP2S_ARCFU	DP2s arcfu
37	44	39.6	577	2	Q89P24	Q89p24 bradyrhizob
38	44	39.6	685	2	Q8IWK5	Q8iwk5 homo sapien
39	44	39.6	784	2	O6LJK5	O6ljk5 photobacter
40	44	39.6	784	2	CAG22525	Cag22525 photobact
41	44	39.6	926	2	O8TE49	O8te49 homo sapien
42	44	39.6	947	2	Q9BIA3	Q9bia3 caenorhabdi
43	44	39.6	1100	2	Q7XR14	Q7xr14 oryza sativ
44	44	39.6	4256	2	Q8MJF3	Q8mjf3 canis famil
45	44	39.6	4311	2	Q7YQK5	Q7yqk5 canis famil

ALIGNMENTS

RESULT 1

Q9HSG0 PRELIMINARY; PRT; 249 AA.
AC Q9HSG0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Vng0250C.
GN OrderedLocusNames=VNG0250.3, VNG0250C;
OS Halobacterium sp. (strain NEC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEC-1 / ATCC 700922 / JCM 11081;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M., Shukla H.D., Lasky S.R., Baiga N.S., Thorsson V., Shrogha J., Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A., Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W., RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H., Ikenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H., Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.; RA InterPro: IPR002855; DUF137.
RT "Genome sequence of Halobacterium species NRC-1";
RL EMBL; AB004988; AAG18847.1; -.
DR PIR; C84185; C84185.
DR InterPro: IPR002855; DUF137.
DR InterPro: IPR011060; RibP_bind_barrel.
DR Pfam; PF02006; DUF137; 1.
DR FIRSF; FIRSF004853; DUF137; 1.
DR ProDom; PD016827; DUF137; 1.
KW Complete proteome.
SQ SEQUENCE 249 AA; 26340 MW; 7957E58363EB88B4 CRC64;

Query Match 41.4%; Score 46; DB 2; Length 249;
Best Local Similarity 63.2%; Pred. No. 34;
Matches 12; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 4 EREWAMFLRAASSRIRGGV 22

Db 229 DREAA-LRAABERIRGGM 245

RESULT 2

Q7DM12 PRELIMINARY; PRT; 251 AA.
AC Q7DM12;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Abrin A-chain (Fragment).

```

OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X76720; CAAS4138.1; -.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGARICIN; 1.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28098 MW; 817B8B5C2A4ABCCE CRC64;

Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 3
Q96235 PRELIMINARY; PRT; 251 AA.
AC Q96235;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Abrin A chain E164A/R167L (Fragment).
GN Name=pcDNAAC-1-E164A/R167L;
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X76721; CAAS4139.1; -.
DR HSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28040 MW; D57FCB182E0EECC9 CRC64;

Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 5
Q96237 PRELIMINARY; PRT; 251 AA.
AC Q96237;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE pcDNAAC-1-E164A/R167L protein (Fragment).
GN Name=pcDNAAC-1-E164A/R167L;
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X76644; CAAS4092.1; -.
DR HSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 27997 MW; 3B60351839AEFE7E CRC64;

```

```

Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 4
Q96236 PRELIMINARY; PRT; 251 AA.
AC Q96236;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Abrin A chain E164A (Fragment).
GN Name=pcDNAAC-1-E164A;
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X76721; CAAS4139.1; -.
DR HSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28040 MW; D57FCB182E0EECC9 CRC64;

Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
DB 11 GATSQSYKQFIEALRRLRGL 32

RESULT 5
Q96237 PRELIMINARY; PRT; 251 AA.
AC Q96237;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE pcDNAAC-1-E164A/R167L protein (Fragment).
GN Name=pcDNAAC-1-E164A/R167L;
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RT derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X76721; CAAS4139.1; -.
DR HSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28040 MW; D57FCB182E0EECC9 CRC64;

```



```

RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RL derived by site-specific mutagenesis in Escherichia coli.";
RL Eur. J. Biochem. 219:83-87(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X54722; CAA54140.1; -.
DR HSSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 251 251
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 6F64755C3DEAFB79 CRC64;

Query Match 41.4%; Score 46; DB 2; Length 251;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGV 22
|| : : | : | : |||
DB 11 GATSQSYKQFIEALRRLRGGL 32

RESULT 6
Q38760 PRELIMINARY; PRT; 252 AA.
AC Q38760;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ricin A-chain type 72 (Fragment).
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaves;
RA Evensen G., Mathiesen A., Sundan A.;
RL Submitted (OCT-1990) to the EMBL/GenBank/DBSJ databases.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X54873; CAA38655.1; -.
DR HSSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 252 252
FT NON_TER 252 252
SQ SEQUENCE 252 AA; 187B8B4E134AEC5 CRC64;

Query Match 41.4%; Score 46; DB 2; Length 252;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGV 22
|| : : | : | : |||
DB 12 GATSQSYKQFIEALRRLRGGL 33

RESULT 8
Q38765 PRELIMINARY; PRT; 319 AA.
AC Q38765;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS OrderedLocustNames=PP3706;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Haden R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Madupu R., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzaz A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohnsels J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Dueterhoeft A., Tuenmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the

```

```

DB 12 GATSQSYKQFIEALRRLRGGL 33

RESULT 7
Q38761 PRELIMINARY; PRT; 252 AA.
ID Q38761 Q96234;
AC Q38761;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ricin A-chain type 73 (Fragment).
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaves;
RA Evensen G., Mathiesen A., Sundan A.;
RL Submitted (OCT-1990) to the EMBL/GenBank/DBSJ databases.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.
DR EMBL; X54873; CAA38655.1; -.
DR HSSP; P11140; IABR.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0030598; F:rRNA N-glycosylase activity; IEA.
DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Plant defense; Protein synthesis inhibitor; Toxin.
FT NON_TER 252 252
FT NON_TER 252 252
SQ SEQUENCE 252 AA; 187B8B4E134AEC5 CRC64;

Query Match 41.4%; Score 46; DB 2; Length 252;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGV 22
|| : : | : | : |||
DB 12 GATSQSYKQFIEALRRLRGGL 33

RESULT 8
Q38765 PRELIMINARY; PRT; 319 AA.
ID Q38765;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS OrderedLocustNames=PP3706;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Haden R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Madupu R., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzaz A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohnsels J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Dueterhoeft A., Tuenmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the

```

```

RT metabolically versatile Pseudomonas putida KT2440.;
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016788; AAN69303.1; -.
DR TIGR; PP3706; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 319 AA; 35856 MW; ACA25B56B8D47F39 CRC64;

Query Match 41.4%; Score 45; DB 2; Length 319;
Best Local Similarity 44.4%; Pred. No. 43;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 MEREWAMEFLRAASSRIG 20
Db 90 IEAWESEFLKSLYSLDGL 107

RESULT 9
Q8VYD3 PRELIMINARY; PRT; 382 AA.
AC Q8VYD3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein At2g46200.
GN Name=At2g46200;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carrinci P., Chen H., Cheuk R., Hayaishizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carrinci P., Chen H., Cheuk R., Hayaishizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY072162; AAL59984.1; -.
DR EMBL; AY096418; AAM20058.1; -.
KW Hypothetical protein.
SQ SEQUENCE 382 AA; 44030 MW; 5BFFC824E1CF7805 CRC64;

Query Match 41.4%; Score 45; DB 2; Length 382;
Best Local Similarity 50.0%; Pred. No. 52;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 339 IRREWDVLSAGGSRIIP 356
Db 339 IRREWDVLSAGGSRIIP 356

RESULT 10
ABRA ABRRP STANDARD; PRT; 528 AA.
AC P11140; P28589;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Abrin-a precursor [Contains: Abrin-a chain (EC 3.2.2.22) (rRNA N-
```

```

DE glycosidase); Abrin-a B chain].
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93132198; PubMed=8421313;
RA Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.;
RT "Primary structure of three distinct isoforms determined by cDNA
RT sequencing. Conservation and significance.";
RL J. Mol. Biol. 229:263-267(1993).
RN [2]
RP SEQUENCE OF 1-251.
RC TISSUE=Seed;
RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;
RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic
RT protein from the seeds of Abrus precatorius.";
RL Agric. Biol. Chem. 52:1095-1097(1988).
RN [3]
RP SEQUENCE OF 1-251 FROM N.A.
RC TISSUE=Leaf;
RA MEDLINE=91201329; PubMed=2016300;
RA Evensen G., Mathiesen A., Sundan A.;
RT "Direct molecular cloning and expression of two distinct abrin A-
RT chains.";
RL J. Biol. Chem. 266:6848-6852(1991).
RN [4]
RP SEQUENCE OF 262-528.
RA MEDLINE=92371656; PubMed=1505674;
RA Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;
RT "The complete primary structure of abrin-a B chain.";
RL PDBS Lett. 309:115-118(1992).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).
RA MEDLINE=95333188; PubMed=7608980;
RA Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
RT "Crystal structure of abrin-a at 2.14 A.";
RL J. Mol. Biol. 250:354-367(1995).
CC -1- FUNCTION: The A chain is responsible for inhibiting protein
CC synthesis through the catalytic inactivation of 60S ribosomal
CC subunits by removing adenine from position 4,324 of 28S rRNA.
CC Abrin-a is more toxic than ricin.
CC -1- FUNCTION: The B chain is a galactose-specific lectin that
CC facilitates the binding of abrin to the cell membrane that
CC precedes endocytosis.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: Disulfide-linked dimer of A and B chains.
CC -1- DOMAIN: The B chain is composed of two domains, each domain
CC consists of 3 homologous subdomains (alpha, beta, gamma).
CC -1- SIMILARITY: In the N-terminal section; belongs to the ribosome-
CC inactivating protein family. Type 2 RIP subfamily.
CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M98344; AAA32624.1; -
CC EMBL; X54872; -; NOT_ANNOTATED_CDS.
CC PIR; S3429; TZLSA.
CC PDB; 1ABR; X-ray; A=1-251, B=262-528.
CC InterPro; IPR008997; RicinB like.
CC InterPro; IPR000772; Ricin_B_lectin.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00652; Ricin_B_lectin; 6.
CC Pfam; PF00161; RIP; 1.

```



```

RC STRAIN-ATCC 15944;
RA Liu W., Nonaka K., Nie L., Bae J., Zazopoulos E., Farnet C.M.,
RL Shen B.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15944;
RX MEDLINE=22447897; PubMed=12536216;
RA Zazopoulos E., Huang K., Staiffa A., Liu W., Bachmann B.O., Nonaka K.,
RA Ahlert J., Thorson J.S., Shen B., Farnet C.M.;
RT "A genomics-guided approach for discovering and expressing cryptic
RT metabolic pathways.";
RL Nat. Biotechnol. 21:187-190(2003).
DR EMBL; AY117439; AAM78015.1; -.
DR EMBL; AF546157; AAO25901.1; -.
KW Hypothetical protein.
SQ SEQUENCE 328 AA; 35711 MW; 7C096FF3C4BA297E CRC64;
Query Match 41.0%; Score 45.5; DB 2; Length 328;
Best Local Similarity 32.4%; Pred. No. 54;
Matches 12; Conservative 5; Mismatches 5; Indels 15; Gaps 1;

QY 1 GAMEREWAMFLR-----AASGRIRGV 22
Db 229 GATEQELALFARQAGEHRWALAQGAFAAEARVRAGL 265

RESULT 12
Q9KD28 PRELIMINARY; PRT; 231 AA.
ID Q9KD28;
AC Q9KD28;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BH1391 protein.
GN Name=BH1391;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001511; BAB05110.1; -.
DR PIR; G83823; G83823.
DR InterPro; IPR008258; SLT.
DR Pfam; PF01464; SLT; 1.
SQ SEQUENCE 231 AA; 27050 MW; F30860E6ACF83ECC CRC64;
Query Match 40.5%; Score 45; DB 2; Length 231;
Best Local Similarity 47.1%; Pred. No. 46;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASR 17
Db 83 GAFEXGNGWFLAGISQ 99

RESULT 13
Q7D052 PRELIMINARY; PRT; 266 AA.
ID Q7D052;
AC Q7D052;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGR_C_1872P.
GN OrderedLocusNames=AGR_C_1872;

```

```

OS Agrobacterium tumefaciens (strain C58 / ATCC 33970);
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
CX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cerson;
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Houmlo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
RA Roumel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Planagan C., Crowell C., Garson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE008032; AAK86824.1; -.
DR InterPro; IPR005000; Hpch_Hpai.
DR Pfam; PF03328; Hpch_Hpai; 1.
SQ SEQUENCE 266 AA; 27934 MW; D37F0891D97D1599 CRC64;
Query Match 40.5%; Score 45; DB 2; Length 266;
Best Local Similarity 52.4%; Pred. No. 53;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASRIRCG 21
Db 236 GADVTEFANTLRALSARYKGG 256

RESULT 14
Q8UGM2 PRELIMINARY; PRT; 266 AA.
ID Q8UGM2;
AC Q8UGM2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Aldolase.
GN OrderedLocusNames=Atul014;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
CX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
DR EMBL; AE009066; AAL42028.1; -.
DR PIR; AF2701; AF2701.
DR PIR; G97483; G97483.
DR HSP; P33522; IDXE.
DR GO; GO:0016228; F:aldolase activity; IEA.
DR InterPro; IPR005000; Hpch_Hpai.
DR Pfam; PF03328; Hpch_Hpai; 1.
KW Complete proteome.
SQ SEQUENCE 266 AA; 27934 MW; D37F0891D97D1599 CRC64;
Query Match 40.5%; Score 45; DB 2; Length 266;
Best Local Similarity 52.4%; Pred. No. 53;

```

Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGG 21
 |||||
 Db 236 GADVTEPANTLRALSARYKGG 256
 |||||

RESULT 15

Q6EVP5
 ID Q6EVP5 PRELIMINARY; PRT; 1423 AA.
 AC Q6EVP5;
 DT 01-OCT-2004 (TRENBLrel. 28, Created)
 DT 01-OCT-2004 (TRENBLrel. 28, Last sequence update)
 DT 01-OCT-2004 (TRENBLrel. 28, Last annotation update)
 DE Putative membrane-bound sugar-binding protein.
 GN Name=yapi89;
 OS Yersinia pseudotuberculosis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=633;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=32777;
 RA Collin F., Billault A., Mallet C., Simonet M., Marceau M.;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ627388; CAF28563.1; -;
 SQ SEQUENCE 1423 AA; 160398 MW; F58C514D7ED8561A CRC64;

Query Match 40.5%; Score 45; DB 2; Length 1423;
 Best Local Similarity 36.4%; Pred. No. 2.8e+02;
 Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAMEREWAMFLRAASSRIRGGV 22
 |||||
 Db 283 GVLGRGNLFWESALSRVDDGI 304
 |||||

Search completed: November 10, 2004, 14:50:10
 Job time : 31.3057 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:44:14 ; Search time 6.58868 Seconds
(without alignments)
181.178 Million cell updates/sec

Title: US-10-092-750-63

Perfect score: 102

Sequence: 1 RGLWDRVLEWGLEPRQ 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50.5	49.5	277	4	US-09-583-110-5300
2	48	47.1	245	4	US-09-252-991A-20982
3	45	44.1	292	4	US-09-710-279-1434
4	45	44.1	297	3	US-09-134-001C-3124
5	45	44.1	431	4	US-09-252-991A-26751
6	45	44.1	580	3	US-08-818-112-75
7	45	44.1	580	3	US-08-818-111-76
8	45	44.1	580	4	US-09-056-556-75
9	45	44.1	580	4	US-09-072-596-76
10	45	44.1	580	4	US-09-072-967-75
11	44	43.1	371	4	US-09-252-991A-21750
12	44	43.1	398	4	US-09-489-039A-7634
13	44	43.1	544	4	US-09-540-236-2905
14	43	42.2	59	3	US-08-840-767-14
15	43	42.2	521	4	US-09-543-681A-5969
16	42.5	41.7	372	4	US-09-252-991A-31788
17	42.5	41.7	510	4	US-09-252-991A-29467
18	42	41.2	111	3	US-08-881-037-20
19	42	41.2	217	3	US-09-248-588-9
20	42	41.2	542	4	US-09-252-991A-32139
21	42	41.2	1051	4	US-09-252-991A-16989
22	42	41.2	1053	4	US-09-543-681A-5981
23	42	41.2	1055	4	US-09-489-039A-9150
24	42	41.2	1503	4	US-09-792-616-3
25	42	41.2	1874	4	US-09-602-787A-46
26	41.5	40.7	357	4	US-09-489-039A-14100
27	41	40.2	257	3	US-08-728-603-19

28 41 40.2 285 4 US-09-489-039A-12915 Sequence 12915, A
29 41 40.2 364 4 US-09-328-352-4956 Sequence 4956, Ap
30 41 40.2 490 4 US-09-252-991A-27210 Sequence 27210, A
31 41 40.2 569 4 US-10-106-275-2 Sequence 2, Appl
32 41 40.2 693 4 US-09-489-039A-8763 Sequence 8763, Ap
33 41 40.2 913 4 US-09-602-787A-170 Sequence 170, Appl
34 41 40.2 1841 2 US-08-804-227C-6 Sequence 6, Appl
35 41 40.2 1891 2 US-08-804-227C-12 Sequence 12, Appl
36 41 40.2 1891 2 US-08-804-198-6 Sequence 6, Appl
37 41 40.2 2818 3 US-09-413-814-28 Sequence 28, Appl
38 41 40.2 4630 4 US-09-091-608-2 Sequence 2, Appl
39 41 40.2 5215 3 US-09-105-537-2 Sequence 2, Appl
40 40.5 39.7 203 4 US-09-489-039A-11397 Sequence 11397, A
41 40.5 39.7 248 3 US-09-134-001C-5271 Sequence 5271, Ap
42 40.5 39.7 318 4 US-09-710-279-2800 Sequence 2800, Ap
43 40.5 39.7 429 1 US-08-218-943-3 Sequence 3, Appl
44 40.5 39.7 458 4 US-09-489-039A-11632 Sequence 11632, A
45 40.5 39.7 459 4 US-09-543-681A-5540 Sequence 5540, Ap

ALIGNMENTS

RESULT 1

US-09-583-110-5300
; Sequence 5300, Application US/09583110
; Patent No. 6699703

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.

; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

; FILE REFERENCE: PATH00-07A

; CURRENT APPLICATION NUMBER: US/09/583,110

; CURRENT FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/107,433

; PRIOR FILING DATE: 1998-06-30

; PRIOR APPLICATION NUMBER: US 60/085,131

; PRIOR FILING DATE: 1998-05-12

; PRIOR APPLICATION NUMBER: US 60/051,553

; PRIOR FILING DATE: 1997-07-02

; NUMBER OF SEQ ID NOS: 5322

; SEQ ID NO 5300

; LENGTH: 277

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-09-583-110-5300

Query Match 49.5%; Score 50.5; DB 4; Length 277;

Best Local Similarity 56.5%; Pred. No. 4.1;

Matches 13; Conservative 3; Mismatches 2; Indels 5; Gaps 2;

Qy 1 RGLWDRVLE---EWGLE-PRQ 18

Db 63 RGNWDRVLEVDGEGYGLPEKE 85

RESULT 2

US-09-252-991A-20982

; Sequence 20982, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 20982

```
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20982

Query Match          47.1%; Score 48; DB 4; Length 245;
Best Local Similarity 53.8%; Pred. No. 8.6;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      6 DVLVEWGLEPRQ 18
DB      50 DRAAEWGLQPE 62
      ||| ||| ||| |||
      ||| ||| ||| |||

RESULT 3
US-09-710-279-1434
; Sequence 1434, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1434
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-279-1434

Query Match          44.1%; Score 45; DB 4; Length 292;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      5 VDRVLEWGLEPRQ 18
DB      165 LDRFIDEYGLNPQ 178
      ||| ||| ||| |||
      ||| ||| ||| |||

RESULT 4
US-09-134-001C-3124
; Sequence 3124, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3124
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3124

Query Match          44.1%; Score 45; DB 3; Length 297;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      5 VDRVLEWGLEPRQ 18
      ||| ||| ||| |||
      ||| ||| ||| |||

us-10-092-750-63.ra1

Db      170 LDRFIDEYGLNPQ 183
      ||| ||| ||| |||
      ||| ||| ||| |||

RESULT 5
US-09-252-991A-26751
; Sequence 26751, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26751
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26751

Query Match          44.1%; Score 45; DB 4; Length 431;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      3 LWDVRLVEEWGL 14
DB      380 VWVSGVIDQWGL 391
      ||| ||| ||| |||
      ||| ||| ||| |||

RESULT 6
US-08-818-112-75
; Sequence 75, Application US/08818112
; Patent No. 6290969
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,112
; FILING DATE: 13-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.411C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
```


INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 580 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-818-112-75

Query Match 44.1%; Score 45; DB 3; Length 580;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGLWDRVLEWGLEPR 17
DB 82 RTLGVRTLSQWNLSPR 98

RESULT 7

US-08-818-111-76
Sequence 76, Application US/08818111
Patent No. 6338852

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonia
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 148
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,111
FILING DATE: 13-MAR-1997
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C6
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 622-6031
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 580 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-818-111-76

Query Match 44.1%; Score 45; DB 3; Length 580;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGLWDRVLEWGLEPR 17
DB 82 RTLGVRTLSQWNLSPR 98

RESULT 8

US-09-056-556-75

Sequence 75, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND

NUMBER OF SEQUENCES: 241

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/056,556

FILING DATE: 07-APR-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Maki, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 210121.457

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 75:

SEQUENCE CHARACTERISTICS:

LENGTH: 580 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-056-556-75

Query Match 44.1%; Score 45; DB 3; Length 580;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGLWDRVLEWGLEPR 17

DB 82 RTLGVRTLSQWNLSPR 98

RESULT 9

US-09-072-596-76

Sequence 76, Application US/09072596
Patent No. 6458366
GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Dillon, Davin C.

APPLICANT: Campos-Neto, Antonia

APPLICANT: Houghton, Raymond

APPLICANT: Vedvick, Thomas S.

APPLICANT: Twardzik, Daniel R.

APPLICANT: Lodes, Michael J.

APPLICANT: Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF

NUMBER OF SEQUENCES: 350

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

TUBERCULOSIS

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: PatentIn Release #1.0, Version #1.30
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 580 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-072-596-76

Query Match 44.1%; Score 45; DB 4; Length 580;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGLWDRVLEEWGLEPR 17
Db 82 RTLGVRTLSQWNLSPR 98

RESULT 10
US-09-072-967-75
Sequence 75, Application US/09072967
Patent No. 6592877
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Twardzik, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 355
CORRESPONDENCE ADDRESS:
ADDRESSES: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: PatentIn Release #1.0, Version #1.30
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:

LENGTH: 580 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-072-967-75

Query Match 44.1%; Score 45; DB 4; Length 580;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 RGLWDRVLEEWGLEPR 17
Db 82 RTLGVRTLSQWNLSPR 98

RESULT 11
US-09-252-991A-21750
Sequence 21750, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21750
LENGTH: 371
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21750

Query Match 43.1%; Score 44; DB 4; Length 371;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 GLWDRVLEEWGLE 15
Db 208 GLGLDKLLRDGWGE 221

RESULT 12
US-09-489-039A-7634
Sequence 7634, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Braton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7634
LENGTH: 398
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7634

Query Match 43.1%; Score 44; DB 4; Length 398;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 LWDRVLEEWGLEPRQ 18
Db 151 LWVERTEBEKINPR 166

```

; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5969
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5969

Query Match      42.2%;   Score 43;   DB 4;   Length 521;
Best Local Similarity 50.0%;   Pred. No. 1.2e+02;
Matches 7;   Conservative 2;   Mismatches 5;   Indels

Qy      1 RGLWDRVLEEKGL 14
          | : |||||
          63 RRFWLPMLLAEWGL 76

Db

Search completed: November 10, 2004, 14:55:44
Job time : 6.63868 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 01:28:30 ; Search time 20.8189 Seconds
(without alignments)
305.399 Million cell updates/sec

Title: US-10-092-750-63

Perfect score: 102
Sequence: 1 RGLWDRVLEEWGLEPRQ 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US10G_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US10H_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	18	US-10-092-750-63	Sequence 63, Appl
2	52	51.0	198	US-10-425-115-244728	Sequence 244728, A
3	52	51.0	3970	US-10-156-761-10429	Sequence 10429, A
4	51	50.0	582	US-10-437-963-147954	Sequence 147954, A
5	50	49.0	174	US-10-767-701-34655	Sequence 34655, A
6	49	48.0	277	US-09-769-787-115	Sequence 115, Appl
7	47	46.1	118	US-10-094-749-2330	Sequence 2330, Ap
8	46	45.1	306	US-10-369-493-9676	Sequence 9676, Ap
9	45.5	44.6	94	US-09-738-626-5765	Sequence 5765, Ap
10	45.5	44.6	161	US-10-108-260A-3255	Sequence 3255, Ap
11	45.5	44.6	176	US-10-002-631C-130	Sequence 130, Appl
12	45.5	44.6	177	US-10-108-260A-3256	Sequence 3256, Ap
13	45	44.1	49	US-10-424-599-148079	Sequence 148079,

14	44.1	131	15	US-10-424-599-267456	Sequence 267456,
15	44.1	290	15	US-10-282-122A-77243	Sequence 77243, A
16	44.1	354	12	US-09-855-604-863	Sequence 863, App
17	44.1	380	14	US-10-369-493-1093	Sequence 1093, Ap
18	44.1	580	14	US-10-193-002-76	Sequence 76, Appl
19	44.1	580	14	US-10-084-843-75	Sequence 75, Appl
20	44.1	602	12	US-09-855-604-886	Sequence 886, App
21	44.1	1059	15	US-10-282-122A-50840	Sequence 50840, A
22	44.5	152	16	US-10-437-963-176611	Sequence 176611,
23	44.5	690	14	US-10-156-761-14817	Sequence 14817, A
24	44	68	15	US-10-424-599-248057	Sequence 248057,
25	44	113	17	US-10-425-115-354403	Sequence 354403,
26	44	220	15	US-10-425-115-45257	Sequence 45257, A
27	44	310	15	US-10-389-647-548	Sequence 548, App
28	44	487	15	US-10-424-599-228961	Sequence 228961,
29	44	578	15	US-10-425-114-70225	Sequence 70225, A
30	44	593	17	US-10-425-115-227309	Sequence 227309,
31	43.5	196	14	US-10-369-493-8559	Sequence 8559, Ap
32	43.5	686	14	US-10-369-493-6778	Sequence 6778, Ap
33	43	73	15	US-10-424-599-190284	Sequence 190284,
34	43	109	9	US-09-205-658-236	Sequence 236, App
35	43	109	10	US-09-963-693-236	Sequence 41102, A
36	43	131	16	US-10-767-701-41102	Sequence 226050,
37	43	153	15	US-10-424-599-236050	Sequence 652, App
38	43	220	14	US-10-286-264-2	Sequence 922, App
39	43	220	14	US-10-228-066A-652	Sequence 286732,
40	43	220	15	US-10-374-780A-2658	Sequence 286735,
41	43	220	15	US-10-413-699B-922	Sequence 43310, A
42	43	258	17	US-10-425-115-286735	Sequence 64659, A
43	43	314	17	US-10-425-115-286735	
44	43	320	15	US-10-425-114-43310	
45	43	320	15	US-10-425-114-64659	

ALIGNMENTS

RESULT 1
US-10-092-750-63
; Sequence 63, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-63

Query Match 100.0%; Score 102; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.5e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGLWDRVLEEWGLEPRQ 18
|||||
Db 1 RGLWDRVLEEWGLEPRQ 18

RESULT 2
US-10-425-115-244728
; Sequence 244728, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 244728
 ; LENGTH: 198
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(198)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_154781C.1.pep
 US-10-425-115-244728

Query Match 51.0%; Score 52; DB 17; Length 198;
 Best Local Similarity 46.7%; Pred. No. 16;
 Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 4 WVDVRLVEWGLEPRQ 18
 DB 76 WFDRIKEFGIEPAE 90

RESULT 3
 US-10-156-761-10429
 ; Sequence 10429, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; PRIOR FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 10429
 ; LENGTH: 3970
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-10429

Query Match 51.0%; Score 52; DB 14; Length 3970;
 Best Local Similarity 56.2%; Pred. No. 3.1e-02;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 GLWVDRVLEWGLEPR 17
 DB 662 GVALFRLLEWGVRRP 677

RESULT 4
 US-10-437-963-147954
 ; Sequence 147954, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 147954
 ; LENGTH: 582
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_48433C.1.pep
 US-10-437-963-147954

Query Match 50.0%; Score 51; DB 16; Length 582;
 Best Local Similarity 53.3%; Pred. No. 63;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 GLWVDRVLEWGLEP 16
 DB 143 GRWVHATLRRWGVPE 157

RESULT 5
 US-10-767-701-34655
 ; Sequence 34655, Application US/10767701
 ; Publication No. US20040172694A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 ; FILE REFERENCE: 38-21(53535)B
 ; CURRENT APPLICATION NUMBER: US/10/767,701
 ; CURRENT FILING DATE: 2004-01-29
 ; NUMBER OF SEQ ID NOS: 63128
 ; SEQ ID NO 34655
 ; LENGTH: 174
 ; TYPE: PRT
 ; ORGANISM: Sorghum bicolor
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(174)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CS1561_1.pep
 US-10-767-701-34655

Query Match 49.0%; Score 50; DB 16; Length 174;
 Best Local Similarity 46.7%; Pred. No. 27;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 WVDVRLVEWGLEPRQ 18
 DB 120 WFDRIKEFGIEPTE 134

RESULT 6
 US-09-769-787-115
 ; Sequence 115, Application US/09769787
 ; Publication No. US20030091577A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Microbial Technics Limited
 ; APPLICANT: Gilbert, Christophe FG
 ; APPLICANT: Hansbro, Philip M

```

; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 115
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-115

Query Match      48.0%; Score 49; DB 10; Length 277;
Best Local Similarity 54.5%; Pred. No. 59;
Matches 12; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 1 RGLWVDRVLE----EWGLEPRQ 18
DB 63 RGNWDRVLEALDGGYGLDPQ 84

RESULT 7
US-10-094-749-2330
; Sequence 2330, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROFUKU
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2330
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2330

Query Match      46.1%; Score 47; DB 14; Length 118;
Best Local Similarity 46.7%; Pred. No. 49;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 WDRVLEWGLEPRQ 18
DB 54 WVPRLQEWKPRQ 68

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9676
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Desulfitobacterium hafniense
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(306)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-9676

Query Match      45.1%; Score 46; DB 14; Length 306;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 1 RGLW--VDRVLEWVG 13
DB 180 RGLWDYIDRVQRQY 194

RESULT 9
US-09-738-626-5765
; Sequence 5765, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5765
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5765

Query Match      44.6%; Score 45.5; DB 9; Length 94;
Best Local Similarity 69.2%; Pred. No. 63;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
```

```

; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3256
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3256

Query Match      44.6%; Score 45.5; DB 15; Length 177;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

Qy 1 RGLWDR-VLE--EWG 13
Db 48 RGLWIRKVLERSWG 63

RESULT 13
US-10-424-599-148079
; Sequence 148079, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 148079
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(49)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_104737C.1.pep
US-10-424-599-148079

Query Match      44.1%; Score 45; DB 15; Length 49;
Best Local Similarity 47.4%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

Qy 2 GLWVDR----VLEEWGLEP 16
Db 16 GLWLSRWFTVVMYWG LHP 34

RESULT 14
US-10-424-599-267456
; Sequence 267456, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 267456
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:

```

```

Qy 5 VDRVLEEWGLEPR 17
Db 75 VDTVIEQWG-BPR 86

US-10-108-260A-3255
; Sequence 3255, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3255
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3255

Query Match      44.6%; Score 45.5; DB 15; Length 161;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 12; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

Qy 1 RGLWDR-VLE--EWG 13
Db 32 RGLWIRKVLERSWG 47

RESULT 11
US-10-002-631C-130
; Sequence 130, Application US/10002631C
; Publication No. US20030157486A1
; GENERAL INFORMATION:
; APPLICANT: Graff, Jonathon M.
; APPLICANT: Muenster, Matthew
; TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
; FILE REFERENCE: A34943 090495.0243
; CURRENT APPLICATION NUMBER: US/10/002,631C
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/300,309
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 130
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-631C-130

Query Match      44.6%; Score 45.5; DB 14; Length 176;
Best Local Similarity 64.3%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

Qy 4 WDRVLEEWGLEPR 17
Db 8 WSDRV---WGAEPR 18

RESULT 12
US-10-108-260A-3256
; Sequence 3256, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27

```



```
; OTHER INFORMATION: Clone ID: PAT_MRT3847_83536C.1.pcp
US-10-424-599-267456
Query Match      44.1%; Score 45; DB 15; Length 131;
Best Local Similarity 47.1%; Pred. No. 1e+02;
Matches      8; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY      3 LWVD--RVLEWGLEPR 17
DB      19 LWSDEPDIWNGLSFR 35
```

```
RESULT 15
US-10-282-122A-77243
; Sequence 77243, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77243
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77243
```

```
Query Match      44.1%; Score 45; DB 15; Length 290;
Best Local Similarity 45.0%; Pred. No. 2.3e+02;
Matches      9; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY      1 RGLW----VDRVLEWGLEP 16
DB      260 RNLWLPAGVDPVKDWGLSP 279
```

Search completed: November 11, 2004, 02:43:19
Job time : 21.8689 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:40:53 ; Search time 4.41509 Seconds
(without alignments)
392.268 Million cell updates/sec

Title: US-10-092-750-63

Perfect score: 102

Sequence: 1 RGLWVDRVLEEWGLEPRQ 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	52.9	1677	2 T18344	P-glycoprotein E -
2	51	50.0	1724	2 T18343	P-glycoprotein - S
3	50	49.0	288	2 G72634	hypothetical prote
4	49	48.0	277	1 B95072	phosphoesterase-re
5	49	48.0	277	1 B97939	phosphoesterase-re
6	47.5	46.6	192	2 S23603	hypothetical prote
7	47	46.1	338	2 B83861	hypothetical prote
8	46	45.1	244	2 T30358	hypothetical prote
9	45	44.1	97	2 T08480	Kleef protein - Ent
10	45	44.1	199	2 T40079	hypothetical prote
11	45	44.1	290	2 D82233	conserved hypothet
12	45	44.1	390	1 B64500	probable hexosyltr
13	45	44.1	437	2 B75067	hypothetical prote
14	45	44.1	1043	2 A80370	multidrug efflux p
15	44.5	43.6	303	2 A45640	phosphoprotein pho
16	44	43.1	97	2 G70603	hypothetical prote
17	44	43.1	245	2 C95314	Tm23b IS ATP-bind
18	44	43.1	245	2 D95334	Tm23b IS ATP-bind
19	44	43.1	265	2 D82566	GumM protein XP236
20	44	43.1	310	2 G83248	probable permease
21	44	43.1	1308	2 T05178	hypothetical prote
22	43.5	42.6	686	2 T15795	hypothetical prote
23	43	42.2	220	2 F84565	probable homeodoma
24	43	42.2	305	2 S42558	phosphoprotein pho
25	43	42.2	305	2 S42559	phosphoprotein pho
26	43	42.2	372	2 S44858	hypothetical prote
27	43	42.2	395	2 B82695	hypothetical prote
28	43	42.2	418	2 S55018	CEM-1 protein - Ca
29	43	42.2	451	2 AG3646	probable tartrate

30 43 42.2 473 2 G72205
31 43 42.2 506 2 T40396
32 43 42.2 518 1 PRBE11
33 43 42.2 715 2 D84549
34 43 42.2 729 2 T04269
35 43 42.2 1249 2 AC1065
36 42.5 41.7 298 2 D71543
37 42.5 41.7 438 2 H83556
38 42 41.2 197 2 T20797
39 42 41.2 162 2 B84152
40 42 41.2 192 1 E84087
41 42 41.2 217 1 NKVLS
42 42 41.2 217 2 H85433
43 42 41.2 224 2 AH1815
44 42 41.2 361 2 A35397
45 42 41.2 389 2 B86634

ALIGNMENTS

RESULT 1

T18344

P-glycoprotein E - Leishmania tropica

C:Species: Leishmania tropica

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T18344

R:Lafuente, B.; Castany, S.; Gamarro, F.

submitted to the EMBL Data Library, April 1996

A:Reference number: Z18880

A:Accession: T18344

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1677 <LAF>

A:Cross-references: UNIPROT:O0805; EMBL:U55381; NID:G1916605; PID:G1916606; PIDN:AA851

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 52.9%; Score 54; DB 2; Length 1677;

Best Local Similarity 56.2%; Pred. No. 9.5;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 RGLWVDRVLEEWGLEP 16

DB 45 RSLNAQRVSDLWGTEP 60

RESULT 2

T18343

P-glycoprotein - Sauroleishmania tarentolae

C:Species: Sauroleishmania tarentolae

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T18343

R:Legare, D.; Hettema, E.; Ouellette, M.

Mol. Biochem. Parasitol. 68, 81-91, 1994

A:Title: The P-glycoprotein-related gene family in Leishmania.

A:Reference number: Z18879; MUID:95198776; PMID:7891750

A:Accession: T18343

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1724 <LEG>

A:Cross-references: UNIPROT:Q25425; EMBL:L29485; NID:G460312; PID:G460313; PIDN:AAA6554

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 50.0%; Score 51; DB 2; Length 1724;

Best Local Similarity 57.1%; Pred. No. 27;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 LWVDRVLEEWGLEP 16

DB 47 LMKQRVIDLWGAEP 60

RESULT 3

A;Map position: 2

Query Match 44.1%; Score 45; DB 2; Length 199;
Best Local Similarity 61.5%; Pred.No. 23;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LWDVRLVEEWGLE 15
||| :||| :|||
Db 59 LMKDHVLRFGLF 71

RESULT 11
D82233
conserved hypothetical protein VC1177 [imported] - Vibrio cholerae (strain N16961 serog C)Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: D82233
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.
Chardonson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,
L.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406 477-483, 2000
A>Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82233
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-290 <HBI>
A:Cross-references: UNIPROT:Q9KSS9; GB:AEO04197; GB:AE003852; NID:g9655642; PIDN:AAF943
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1177
A:Map position: 1
C:Superfamily: hypothetical protein HI1400

Query Match 44.1%; Score 45; DB 2; Length 290;
Best Local Similarity 45.0%; Pred.No. 33;
Matches 9; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

Qy 1 RGLW----VDVRLVEEWGLEP 16
||| |:::|||
Db 260 RNLLPWAGVPVKDWGLSP 279

RESULT 12
F64500
probable hexosyltransferase (EC 2.4.1.-) MJ1607 - Methanococcus jannaschii
N:Alternate names: probable lipopolysaccharide N-acetylglucosaminyltransferase rfbU
C:Species: Methanococcus jannaschii
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: F64500
R:Bult, C.J.; Overbeek, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
Reich, C.I.; Whitehead, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.
rson, J.D.; Sadov, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese,
A>Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:868087
A:Accession: F64500
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-390 <BUL>
A:Cross-references: UNIPROT:Q59002; GB:U67601; GB:L77117; NID:g2826439; PIDN:AAE99629.1
C:Genetics:
A:Map position: FOR1581714-1582886
C:Superfamily: probable hexosyltransferase ytxN
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 44.1%; Score 45; DB 1; Length 390;
Best Local Similarity 37.5%; Pred.No. 45;
Matches 9; Conservative 2; Mismatches 1; Indels 12; Gaps 1;

Qy 2 GLW-----VDVRLVEEWG 13
|::| ||||| :

Db 328 GIWYKPNPDSIANGVDRVLSDMG 351

RESULT 13

B75067
hypothetical protein BAB1355 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: B75067
R:Anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
A:Reference number: A75001
A:Accession: B75067
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-437 <KAW>
A:Cross-references: UNIPROT:Q9UV11; GB:AJ248287; GB:AL096836; NID:G5458657; PIDN:CA35043
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: BAB1355

Query Match 44.1%; Score 45; DB 2; Length 437;

Best Local Similarity 61.5%; Pred. No. 51;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 WVDRLVEEWGLEP 16

DB 31 WVPRIEDISLEP 43

RESULT 14

AB0370
multidrug efflux protein [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB0370
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, N.; Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0370
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1043 <KUR>
A:Cross-references: UNIPROT:Q8ZCE8; GB:AL590842; PIDN:CA32285.1; PID:G15980996; GSPDB:G
C:Genetics:
A:Gene: YPO3043
C:Superfamily: acriflavin resistance protein

Query Match 44.1%; Score 45; DB 2; Length 1043;

Best Local Similarity 57.1%; Pred. No. 1.3e+02;

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 VDRVLEEWGLEPRQ 18

DB 416 VERVMSERGLDPR 429

RESULT 15

A45640
phosphoprotein phosphatase (EC 3.1.3.16) 2A catalytic chain - Trypanosoma brucei
C:Species: Trypanosoma brucei
C:Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 16-Aug-2004
C:Accession: A45640
R:Erondy, N.E.; Donelson, J.E.
Mol. Biochem. Parasitol. 49, 303-314, 1991
A:Title: Characterization of trypanosome protein phosphatase 1 and 2A catalytic subunits
A:Reference number: A45640; MUID:92131067; PMID:1663579
A:Accession: A45640
A:Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-303 <ERO>
A:Cross-references: UNIPROT:Q27787; GB:M74168; NID:G162210; PIDN:AAA73084.1; PID:G162211
A:Experimental source: subsp. rhodesiense
A:Note: sequence extracted from NCBI backbone (NCBIN:78241, NCBIP:78244)
C:Superfamily: Serine/threonine protein phosphatase; phosphoesterase core homology; phospho-
C:Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci-
F:17-276/Domain: phosphoprotein phosphatase homology <PPP>
F:45-113/Domain: phosphoesterase core homology <PEC>
F:51,53,79/Binding site: iron (Asp, His, Asp) #status predicted
F:79,111,161,235/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:82,112,259/Active site: Asp, His, Tyr #status predicted
F:83,208/Binding site: substrate phosphate (Arg) #status predicted

Query Match 43.6%; Score 44.5; DB 2; Length 303;

Best Local Similarity 56.2%; Pred. No. 42;

Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 3 LWVD-RVLEEWGLEPR 17

DB 193 LMSDPFEEEGWGLSPR 208

Search completed: November 10, 2004, 14:52:29

Job time: 5.41509 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 13:38:57 ; Search time 23.9774 Seconds
(without alignments)
431.938 Million cell updates/sec

Title: US-10-092-750-63

Perfect score: 102

Sequence: 1 RGLWDRVLEEWGLPRQ 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt.02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	53.9	241	2 Q72KQ3	Q72KQ3 thermus the
2	55	53.9	241	2 AAS80683	AAS80683 thermus t
3	54	52.9	1677	2 O00805	O00805 leishmania
4	52	51.0	627	2 Q8ZY05	Q8ZY05 pyrobaculum
5	52	51.0	3970	2 Q93H78	Q93H78 streptomyces
6	51	50.0	245	2 Q8GQ42	Q8GQ42 pseudomonas
7	51	50.0	512	2 Q7EX16	Q7EX16 oryza sativ
8	51	50.0	1724	2 BAD10739	BAD10739 oryza sat
9	51	50.0	1724	2 Q25425	Q25425 leishmania
10	50	49.0	288	2 Q9YBR5	Q9YBR5 aeropyrum p
11	50	49.0	293	2 Q7MM50	Q7MM50 vibrio vuln
12	50	49.0	293	2 Q8D8B7	Q8D8B7 vibrio vuln
13	50	49.0	842	2 Q74E16	Q74E16 geobacter s
14	50	49.0	842	2 AAR34273	AAR34273 geobacter
15	49	48.0	277	2 Q8DQ08	Q8DQ08 streptococc
16	49	48.0	277	2 Q97S06	Q97S06 streptococc
17	49	48.0	321	2 Q8P0H0	Q8P0H0 streptococc
18	49	48.0	321	2 Q92Z82	Q92Z82 streptococc
19	49	48.0	321	2 Q7CF02	Q7CF02 streptococc
20	49	48.0	321	2 Q7CN48	Q7CN48 streptococc
21	48	47.1	107	2 Q9WB08	Q9WB08 human immun
22	48	47.1	239	2 Q706M5	Q706M5 pseudomonas
23	48	47.1	239	2 CAE92925	CAE92925 pseudomon
24	48	47.1	362	2 Q6JEI3	Q6JEI3 pichia memb
25	48	47.1	362	2 RAT12523	RAT12523 pichia me
26	48	47.1	373	2 Q6H188	Q6H188 issatchenki
27	48	47.1	379	2 Q8H193	Q8H193 pichia ferm
28	48	47.1	396	2 Q80X17	Q80X17 mus musculu
29	48	47.1	397	2 Q8BLU5	Q8BLU5 mus musculu
30	48	47.1	905	2 Q9UVA0	Q9UVA0 issatchenki
31	48	47.1	956	2 Q93RE7	Q93RE7 bacillus sp

32 47.5 46.6 192 2 Q53897 streptomyces
33 47 46.1 173 2 Q6QH09 alcaligenes
34 47 46.1 173 2 AAS49425 alcaligenes
35 47 46.1 338 2 Q9KC84 bacillus ha
36 47 46.1 388 2 Q8KL87 rhizobium e
37 47 46.1 391 2 Q8A7A2 bacteroides
38 47 46.1 631 2 Q8QJD3 calocera co
39 47 46.1 631 2 AAS57792 calocera
40 46.5 45.6 174 2 Q7Z1M4 trypanosoma
41 46.5 45.6 1151 2 Q6P5L5 chlamydomon
42 46.5 45.6 1151 2 AAS92601 chlamydomon
43 46 45.1 244 2 Q9YMW3 lymantria d
44 46 45.1 323 2 Q8XSC4 ralstonia s
45 46 45.1 363 2 Q8U4S7 pyrobaculum

ALIGNMENTS

RESULT 1

Q72KQ3 PRELIMINARY; PRT; 241 AA.
AC Q72KQ3;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Branched-chain amino acid transport ATP-binding protein livg.
GN Name=livg; OrderedlocusNames=IFC0335;
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=262724;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15084768;
RA Henne A., Brueggemann H., Raasch C., Wierzer A., Hartsch T.,
RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
RA Jacobi C., Starkuviene V., Schlentz S., Dencker S., Huber R.,
RA Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;
RT "The genome sequence of the extreme thermophile Thermus
RT thermophilus".
RL Nat. Biotechnol. 22:547-553(2004).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF017302; AAS80683.1;
DR GO; GO:0005524; F:ATP binding; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 241 AA; 26385 MW; 549D15A22276994D CRC64;

Query Match 53.9%; Score 55; DB 2; Length 241;
Best Local Similarity 78.6%; Pred.No. 4.5;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 WVDRLVEEWGLPR 17

|||||
119 WDRVLELTGLAPR 132

RESULT 2

AAS80683 PRELIMINARY; PRT; 241 AA.
AC AAS80683;
DT 14-APR-2004 (Tremblrel. 27, Created)
DT 14-APR-2004 (Tremblrel. 27, Last sequence update)
DT 11-MAY-2004 (Tremblrel. 27, Last annotation update)
DE Branched-chain amino acid transport ATP-binding protein livg.
GN livg OR TTC0335.
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).

DB 45 RSLWAQRVSLWGTEP 60

RESULT 4

Q8ZY05 PRELIMINARY; PRT; 627 AA.

ID Q8ZY05

AD 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Acylamino-acid-releasing enzyme, conjectural.

GN OrderedLocusNames=PAE1002;

DE Pyrobaculum aerophilum.

OS Archaea; Crenarchasota; Thermoprotei; Thermoproteales;

OC Thermoproteaceae; Pyrobaculum.

OX NCBI_TaxID=13773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=IM2 / ATCC 51768 / DSM 7523;

RX MEDLINE=21664397; PubMed=11792569;

RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I., Miller J.H.;

RA "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum aerophilum";

RT proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).

RL EMBL; AE009798; AAL63191.1; "

DR GO: 0003824; F.catalytic activity; IEA.

DR InterPro; IPR000379; Ser_estrs.

DR InterPro; IPR011042; TolE_C.

DR Complete proteome.

SK SEQUENCE 627 AA; 69906 MW; D115B4D625AF09DF CRC64;

QY Query Match 51.0%; Score 52; DB 2; Length 627;

DB Best Local Similarity 64.7%; Pred. No. 34;

Matches 11; Conservative 0; Mismatches 2; Indels 4; Gaps

QY 2 GLWVDRVLEEWGLEPRQ 18

DB ||||| ||||| |||||

87 GLWV---AEWGGEPRQ 99

RESULT 5

Q93H18 PRELIMINARY; PRT; 3970 AA.

ID Q93H18

AD 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Modular polyketide synthase.

GN Name=olm4; OrderedLocusNames=SAV2892;

OS Streptomyces avermitilis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomycetes.

OX NCBI_TaxID=33903;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MA-4680;

RX MEDLINE=21477403; PubMed=11572948;

RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C., Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T., Kikuchi H., Shiba T., Sakaki Y., Hattori M.;

RA "Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites";

RT proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MA-4680;

RX MEDLINE=22608306; PubMed=12692562;

RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T., Sakaki Y., Hattori M., Omura S.;

RA "Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis";

RT microorganism Streptomyces avermitilis; "


```
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AB070940; BAB69199.1; -.
DR EMBL; AP005032; BAC70603.1; -.
DR HSSP; P25715; 1MLA.
DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0003960; F:NADPH:quinone reductase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR009081; ACP like
DR InterPro; IPR001227; ACP_transferase.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR011032; GroES like.
DR InterPro; IPR000794; ketoacyl synth.
DR InterPro; IPR006162; Pantone S.
DR InterPro; IPR006163; Pp_bind.
DR InterPro; IPR002364; QOR_zeta_crystal.
DR Pfam; PF00698; Acyl_transf_1; 2.
DR Pfam; PF00107; ADH_zinc_N; 1.
DR Pfam; PF00023; Ank; 2.
DR Pfam; PF00109; ketoacyl-synt; 2.
DR Pfam; PF02801; ketoacyl-synt C; 2.
DR Pfam; PF00550; PP-binding; 2.
DR PROSITE; PS00375; ACP_DOMAIN; 2.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
DR PROSITE; PS01162; QOR_ZETA_CRYSTAL; 1.
DR KW Complete proteome; Phosphopantetheine; Transferase.
SQ SEQUENCE 3970 AA; 412662 MW; A93E583FAFA68C47 CRC64;

Query Match 51.0%; Score 52; DB 2; Length 3970;
Best Local Similarity 56.2%; Pred. No. 2.2e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 GLWVDRVLEEWGLEPR 17
DB 662 GVALFRLEEWGVPR 677

RESULT 6
Q8GQ42 PRELIMINARY; PRT; 245 AA.
ID Q8GQ42
AC Q8GQ42;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein ORF C68.
GN Name=ORF C68;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C;
RX MEDLINE=2313472; PubMed=12426355;
RA Larbig K.D., Christmann A., Johann A., Klockgether J., Hartsch T.,
RA Merkl R., Wiehmann L., Fritz H.J., Tummeler B.;
RT "Gene islands integrated into tRNA(Gly) genes confer genome diversity
on a Pseudomonas aeruginosa clone.";
RL J. Bacteriol. 184:6665-6680(2002).
DR EMBL; AF440523; AAN62162.1; -.
KW Hypothetical protein.
SQ SEQUENCE 245 AA; 27314 MW; 86DA12B9069ACA7A CRC64;

Query Match 50.0%; Score 51; DB 2; Length 245;
Best Local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 DRVLEEWGLEPRQ 18

RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AB070940; BAB69199.1; -.
DR EMBL; AP005032; BAC70603.1; -.
DR HSSP; P25715; 1MLA.
DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0003960; F:NADPH:quinone reductase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR009081; ACP like
DR InterPro; IPR001227; ACP_transferase.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR011032; GroES like.
DR InterPro; IPR000794; ketoacyl synth.
DR InterPro; IPR006162; Pantone S.
DR InterPro; IPR006163; Pp_bind.
DR InterPro; IPR002364; QOR_zeta_crystal.
DR Pfam; PF00698; Acyl_transf_1; 2.
DR Pfam; PF00107; ADH_zinc_N; 1.
DR Pfam; PF00023; Ank; 2.
DR Pfam; PF00109; ketoacyl-synt; 2.
DR Pfam; PF02801; ketoacyl-synt C; 2.
DR Pfam; PF00550; PP-binding; 2.
DR PROSITE; PS00375; ACP_DOMAIN; 2.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; 1.
DR PROSITE; PS01162; QOR_ZETA_CRYSTAL; 1.
DR KW Complete proteome; Phosphopantetheine; Transferase.
SQ SEQUENCE 3970 AA; 412662 MW; A93E583FAFA68C47 CRC64;

Query Match 51.0%; Score 52; DB 2; Length 3970;
Best Local Similarity 56.2%; Pred. No. 2.2e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 GLWVDRVLEEWGLEPR 17
DB 662 GVALFRLEEWGVPR 677

RESULT 6
Q8GQ42 PRELIMINARY; PRT; 245 AA.
ID Q8GQ42
AC Q8GQ42;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein ORF C68.
GN Name=ORF C68;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C;
RX MEDLINE=2313472; PubMed=12426355;
RA Larbig K.D., Christmann A., Johann A., Klockgether J., Hartsch T.,
RA Merkl R., Wiehmann L., Fritz H.J., Tummeler B.;
RT "Gene islands integrated into tRNA(Gly) genes confer genome diversity
on a Pseudomonas aeruginosa clone.";
RL J. Bacteriol. 184:6665-6680(2002).
DR EMBL; AF440523; AAN62162.1; -.
KW Hypothetical protein.
SQ SEQUENCE 245 AA; 27314 MW; 86DA12B9069ACA7A CRC64;

Query Match 50.0%; Score 51; DB 2; Length 245;
Best Local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 DRVLEEWGLEPRQ 18

Db 46 DRVATEWGLQPOE 58
||| ||| ||| |||
RESULT 7
Q7EXY6 PRELIMINARY; PRT; 512 AA.
ID Q7EXY6
AC Q7EXY6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Selenium binding protein-like protein.
GN Names=OSJNBa0016N23.128;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006049; BAD10739.1; -.
DR InterPro; IPR002885; PPR.
DR InterPro; IPR008940; Prenyl_trans.
DR Pfam; PF01535; PPR; 7.
DR TIGRFAMs; TIGR00756; PPR; 5.
SQ SEQUENCE 512 AA; 55624 MW; 7B13DCLC3B5FB4D6 CRC64;

Query Match 50.0%; Score 51; DB 2; Length 512;
Best Local Similarity 53.3%; Pred. No. 39;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 GLWVDRVLEEWGLEPR 16
DB 177 GRWVATLRWGVEP 191

RESULT 8
BAD10739 PRELIMINARY; PRT; 512 AA.
ID BAD10739
AC BAD10739;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Selenium binding protein-like protein.
GN OSJNBa0016N23.128.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
clone:OSJNBa0016N23.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006049; BAD10739.1; -.
SQ SEQUENCE 512 AA; 55624 MW; 7B13DCLC3B5FB4D6 CRC64;

Query Match 50.0%; Score 51; DB 2; Length 512;
Best Local Similarity 53.3%; Pred. No. 39;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 GLWVDRVLEEWGLEPR 16
DB 177 GRWVATLRWGVEP 191

RESULT 9
Q25425
```

ID Q25425 PRELIMINARY; PRT; 1724 AA.
 AC Q25425;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE P-glycoprotein.
 OS Leishmania tarentolae (Saurleishmania tarentolae).
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania;
 OC lizard Leishmania.
 OX NCBI_TaxID=5689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TatII;
 RX MEDLINE=95198776; PubMed=7891750;
 RA Legare D., Hettner E., Quellerie M.,
 RT "The P-glycoprotein-related gene family in Leishmania.";
 RL Mol. Biochem. Parasitol. 68:81-91(1994).
 CC -!- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; L29485; AAA65541.1; -;
 DR PIR; T18343; T18343.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0005524; P: ATP binding; IEA.
 DR GO; GO:0004009; P: ATP-binding cassette (ABC) transporter acti. . .; IEA.
 DR GO; GO:0000166; F: nucleotide binding; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR011527; ABC membrane 1.
 DR InterPro; IPR001140; ABC TM transp.
 DR InterPro; IPR003439; ABC transporter.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR SMART; SM00382; AAA; 2.
 DR PROSITE; PS00929; ABC_TM1F; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
 DR ATP-binding.
 KW SEQUENCE 1724 AA; 188967 MW; A8P6CD4A65FA195 CRC64;
 SQ
 Query Match 50.0%; Score 51; DB 2; Length 1724;
 Best Local Similarity 57.1%; Pred. No. 1.3e+02;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 3 LWVDVRLVLEWGLEP 16
 DB 47 LMKQVIDLWGAEP 60
 RESULT 10
 Q9YBR5 PRELIMINARY; PRT; 288 AA.
 ID Q9YBR5
 AC Q9YBR5;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein APE1534.
 GN OrderedLocusNames=APE1534;
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 OC Desulfurococcales; Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 Crenarchaeon, Aeropyrum pernix X1.";
 RL DNA Res. 6:83-101(1999).

DR EMBL; AP000061; BAA80533.1; -;
 DR PIR; G72634; G72634.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 288 AA; 32176 MW; 40235FA12BDD6AFB CRC64;
 Query Match 49.0%; Score 50; DB 2; Length 288;
 Best Local Similarity 58.8%; Pred. No. 31;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 RGLWVDRVLEWGLEP 17
 DB 151 RGLGSDRVVAVWLDLR 167
 RESULT 11
 Q7NM50 PRELIMINARY; PRT; 293 AA.
 ID Q7NM50
 AC Q7NM50;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Predicted metal-dependent phosphoesterase.
 GN Names=VVI223;
 OS Vibrio vulnificus (strain XJ016).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=196600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14656965;
 RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
 RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.-B.-T., Li J.-C., Su T.-L.,
 RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.,
 RT "Comparative genome analysis of Vibrio vulnificus, a marine
 pathogen.";
 RL Genome Res. 13:2577-2587(2003).
 DR EMBL; AP005334; BAC93987.1; -;
 DR GO; GO:0003677; F: DNA binding; IEA.
 DR GO; GO:0003887; F: DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0006280; P: DNA replication; IEA.
 DR InterPro; IPR003141; Pesterase_PHP_N.
 DR InterPro; IPR004013; PHP_C.
 DR Pfam; PF02811; PHP_C; 1.
 DR Pfam; PF02231; PHP_N; 1.
 SQ SEQUENCE 293 AA; 33152 MW; 62C1C56AF17720BC CRC64;
 Query Match 49.0%; Score 50; DB 2; Length 293;
 Best Local Similarity 50.0%; Pred. No. 31;
 Matches 10; Conservative 3; Mismatches 3; Indels 4; Gaps 1;
 QY 1 RGLW----VDRVLEWGLEP 16
 DB 260 RNLALPAGVPEPVKDWGLEP 279
 RESULT 12
 Q8D8B7 PRELIMINARY; PRT; 293 AA.
 ID Q8D8B7
 AC Q8D8B7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Predicted metal-dependent phosphoesterases.
 GN OrderedLocusNames=VVI3063;
 OS Vibrio vulnificus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=672;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=QWC66;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RA Choy H.E.;

RT "Complete genome sequence of *Vibrio vulnificus* CMCP6.";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE016807; AA011387.1; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
 DR GO; GO:0006260; P:DNA replication; IEA.
 DR InterPro; IPR009056; Cytochrome C.
 DR InterPro; IPR003141; Pesterase_PHP_N.
 DR Pfam; PF02811; PHP_C; 1.
 DR Pfam; PF02231; PHP_N; 1.
 DR SMART; SM00481; POLI1IAC; 1.
 DR Complete proteome.
 KW SEQUENCE 293 AA; 33178 MW; 5839B6813C495140 CRC64;

Query Match 49.0%; Score 50; DB 2; Length 293;
 Best Local Similarity 50.0%; Pred. No. 31;
 Matches 10; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

QY 1 RGLW----VDRVLEEWGLEP 16
 ||||| :|:|||||
 Db 260 RNLWLPAGVEPVKDWGLEP 279

RESULT 13

Q74EL6 PRELIMINARY; PRT; 842 AA.
 ID Q74EL6
 AC Q74EL6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Sensory box/GDEF family protein.
 DE ORNames=GSU0946;
 OS Geobacter sulfurreducens.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
 OC Geobacteraceae; Geobacter.
 ON NCBI_TaxID=35554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCA / ATCC 51573;
 RX PubMed=14671304; DOI=10.1126/science.1088727;
 RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
 RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beanan M.J., Dodson R.J.,
 RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
 RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
 RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
 RA Weidman J.F., Kouri H.M., Feldblyum T.V., Utterback T.R.,
 RA Van Aken S.E., Lovley D.R., Fraser C.M.;
 RA "Genome of *Geobacter sulfurreducens*: metal reduction in subsurface environments.";
 RT Science 302:1967-1969(2003).
 RL EMBL; AE017180; AAR34273.1; -;
 DR TIGR; GSU0946; -;
 DR SEQUENCE 842 AA; 92860 MW; 989697FE4637E9AD CRC64;

Query Match 49.0%; Score 50; DB 2; Length 842;
 Best Local Similarity 83.3%; Pred. No. 91;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DR PROSITE; PS50885; HAMP; 1.
 DR PROSITE; PS50113; PAC; 1.
 DR PROSITE; PS50112; PAS; 1.
 SQ SEQUENCE 842 AA; 92860 MW; 989697FE4637E9AD CRC64;

Query Match 49.0%; Score 50; DB 2; Length 842;
 Best Local Similarity 83.3%; Pred. No. 91;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 VDRVLEEWGLEP 16
 ||||| :|:|||||
 Db 680 VDRVLEETGLDP 691

RESULT 14

AAR34273 PRELIMINARY; PRT; 842 AA.
 ID AAR34273
 AC AAR34273;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Sensory box/GDEF family protein.
 DE GSU0946.
 OS Geobacter sulfurreducens.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
 OC Geobacteraceae; Geobacter.
 ON NCBI_TaxID=35554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCA / ATCC 51573;
 RX PubMed=14671304;
 RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
 RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beanan M.J., Dodson R.J.,
 RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
 RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
 RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
 RA Weidman J.F., Kouri H.M., Feldblyum T.V., Utterback T.R.,
 RA Van Aken S.E., Lovley D.R., Fraser C.M.;
 RA "Genome of *Geobacter sulfurreducens*: metal reduction in subsurface environments.";
 RT Science 302:1967-1969(2003).
 RL EMBL; AE017210; AAR34273.1; -;
 DR TIGR; GSU0946; -;
 DR SEQUENCE 842 AA; 92860 MW; 989697FE4637E9AD CRC64;

Query Match 49.0%; Score 50; DB 2; Length 842;
 Best Local Similarity 83.3%; Pred. No. 91;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 VDRVLEEWGLEP 16
 ||||| :|:|||||
 Db 680 VDRVLEETGLDP 691

RESULT 15

Q8DQ08 PRELIMINARY; PRT; 277 AA.
 ID Q8DQ08
 AC Q8DQ08;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein spr0544.
 GN OrderedlocusNames=spr0544;
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 ON NCBI_TaxID=171101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429245; PubMed=11544234;
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,

RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-W., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rosteck P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMEL; AE008433; AAK99348.1; -;
DR PIR; H97939; H97939.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR004843; M:pesterase.
DR InterPro; IPR011152; Pesterase_MJ0912.
DR InterPro; IPR006186; T:phatase_apah.
DR Pfam; PF00149; Metallophos; 1.
DR PIRSF; PIRSF000883; Pesterase_MJ0912; 1.
DR ProDom; PD000252; T:phatase_apah; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 277 AA; 31507 MW; 7A79F94DCA187F9 CRC64;

Query Match 48.0%; Score 49; DB 2; Length 277;
Best Local Similarity 54.5%; Pred. No. 42;
Matches 12; Conservative 2; Mismatches 4; Indels 4; Gaps 1;
OY 1 RGLWDRVLE----EWGLEPRQ 18
DB 63 RGNWDRVLEALDQGYGLEDPQ 84

Search completed: November 10, 2004, 14:50:36
Job time : 25.9774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 12.9537 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-64

Perfect score: 157
Sequence: 1 FVRSGWRLQIGDDMDHAICGHVDRLG 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	52	33.1	614	4	US-09-328-352-5476
2	50	31.8	88	4	US-09-543-681A-5323
3	50	31.8	579	4	US-09-198-452A-918
4	48	30.6	375	3	US-09-323-872A-33
5	48	30.6	375	4	US-09-072-433-18
6	48	30.6	1114	2	US-08-576-626A-31
7	47.5	30.3	4544	1	US-08-469-486-52
8	47.5	30.3	4544	2	US-08-469-486-52
9	47	29.9	358	4	US-09-679-279-16
10	47	29.9	366	4	US-09-328-352-6796
11	47	29.9	465	4	US-09-252-991A-16662
12	47	29.9	934	4	US-09-252-991A-25635
13	46	29.3	27	4	US-09-544-664B-15
14	46	29.3	32	2	US-08-706-741B-55
15	46	29.3	32	2	US-08-924-695A-55
16	46	29.3	33	2	US-08-706-741B-51
17	46	29.3	33	2	US-08-924-695A-53
18	46	29.3	44	2	US-08-706-741B-56
19	46	29.3	44	2	US-08-924-695A-56
20	46	29.3	55	2	US-08-706-741B-52
21	46	29.3	55	2	US-08-924-695A-52
22	46	29.3	122	2	US-08-706-741B-51
23	46	29.3	122	2	US-08-924-695A-51
24	46	29.3	136	3	US-09-136-879-5
25	46	29.3	195	2	US-08-706-741B-6
26	46	29.3	195	2	US-08-924-695A-6
27	46	29.3	195	3	US-09-136-879-2

28	46	29.3	273	4	US-09-270-767-43062	Sequence 43062, A
29	46	29.3	581	3	US-08-961-083-132	Sequence 132, App
30	46	29.3	581	4	US-09-536-784-132	Sequence 132, App
31	46	29.3	623	4	US-09-252-991A-22906	Sequence 22906, A
32	46	29.3	737	4	US-09-583-110-5075	Sequence 5075, Ap
33	46	29.3	783	2	US-08-922-837-2	Sequence 2, Appli
34	46	29.3	783	3	US-09-351-550-2	Sequence 2, Appli
35	45.5	29.0	388	4	US-09-252-991A-30608	Sequence 30608, A
36	45.5	29.0	482	4	US-09-252-991A-24228	Sequence 24228, A
37	45	28.7	488	4	US-09-252-991A-23079	Sequence 23079, A
38	45	28.7	581	4	US-09-252-991A-25414	Sequence 25414, A
39	45	28.7	1036	4	US-09-489-039A-9371	Sequence 9371, Ap
40	45	28.7	1489	6	5183745-2	Patent No. 5183745
41	45	28.7	1706	3	US-08-669-785-2	Sequence 2, Appli
42	45	28.7	1794	6	5183745-6	Patent No. 5183745
43	44	28.0	20	2	US-08-706-741B-85	Sequence 85, Appl
44	44	28.0	20	2	US-08-924-695A-85	Sequence 85, Appl
45	44	28.0	32	2	US-08-706-741B-87	Sequence 87, Appl

ALIGNMENTS

RESULT 1
US-09-328-352-5476
; Sequence 5476, Application US/09328352
; Patent No. 6562958

; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5476
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5476

Query Match 33.1%; Score 52; DB 4; Length 614;
Best Local Similarity 40.9%; Pred. No. 9;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 6 GWRQLQIGDDMDHAICGHVDRL 27
DB 289 GWKIETGVDDIAWIKPGEDRL 310
|||:::|||||

RESULT 2
US-09-543-681A-5323
; Sequence 5323, Application US/09543681A
; Patent No. 6605709

; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5323
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5323

Query Match 31.8%; Score 50; DB 4; Length 88;
Best Local Similarity 55.6%; Pred. No. 2;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 FVRSVGVRLQNGDDMDH 18
DB 71 FRTVGVCPANIGDHADH 88

RESULT 3
US-09-198-452A-918
; Sequence 318, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198.452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 918
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-918

Query Match 31.8%; Score 50; DB 4; Length 579;
Best Local Similarity 45.5%; Pred. No. 17;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 6 GRLQNLGDDMDHAIICGHDVRL 27
DB 260 GWKIEICIGDIAWIRPGRDGL 281

RESULT 4
US-09-323-872A-33
; Sequence 33, Application US/09323872A
; Patent No. 6395539
; GENERAL INFORMATION:
; APPLICANT: Coschigano, Peter
; TITLE OF INVENTION: Compositions and Methods for Bioremediation
; FILE REFERENCE: OHU-03640
; CURRENT APPLICATION NUMBER: US/09/323.872A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 09/072.433
; PRIOR FILING DATE: 1998-05-04
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Thauera aromatica
US-09-323-872A-33

Query Match 30.6%; Score 48; DB 3; Length 375;
Best Local Similarity 52.9%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 SVGVRLQNGDDMDHAI 20
DB 244 TVGWPLQPLDNIHLEHI 260

RESULT 5
US-09-072-433-18
; Sequence 18, Application US/09072433
; Patent No. 6551814
; GENERAL INFORMATION:
; APPLICANT: Coschigano, Peter W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR BIOREMEDIATION
; TITLE OF INVENTION: BIOREMEDIATION
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072.433
FILING DATE: 04-May-1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OHU-03344
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-072-433-18

Query Match 30.6%; Score 48; DB 4; Length 375;
Best Local Similarity 52.9%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 SVGVRLQNGDDMDHAI 20
DB 244 TVGWPLQPLDNIHLEHI 260

RESULT 6
US-08-576-626A-31
; Sequence 31, Application US/08576626A
; Patent No. 5998194
; GENERAL INFORMATION:
; APPLICANT: Summers, R.G.
; APPLICANT: Katz, L. L.
; APPLICANT: Donadio, S.
; APPLICANT: Staver, M.J.
; TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR
; TITLE OF INVENTION: BIOSYNTHESIS GENES
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: Illinois
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/576.626A
FILING DATE: 21-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dianne Casuto

REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 5857.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (847) 938-3137
TELEFAX: (847) 938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 1114 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5998194e
US-08-576-626A-31

Query Match 30.6%; Score 48; DB 2; Length 1114;
Best Local Similarity 34.8%; Pred. No. 75;
Matches 8; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 2 VRSVGWRLQNGDDMDHAICGH 24
DB 16 IRLHLWGYGNGDPYPMLLCGHD 38

RESULT 7
US-08-469-486-52
Sequence 52, Application US/08469486
Patent No. 5739281

GENERAL INFORMATION:
APPLICANT: Thøgersen, Hans Christian
APPLICANT: Holtet, Thor Las
APPLICANT: Etzerodt, Michael
TITLE OF INVENTION: Improved method for the refolding of
TITLE OF INVENTION: Proteins
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,486
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/192,060
FILING DATE: February 4, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Paul T. Clark
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06363/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 542 5070
TELEFAX: 617 542 8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 4544 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-486-52

Query Match 30.3%; Score 47.5; DB 1; Length 4544;

Best Local Similarity 39.1%; Pred. No. 4.5e-02;
Matches 9; Conservative 3; Mismatches 8; Indels 3; Gaps 1;
QY 4 SVGWRLQNGDDMDHAICGH 23
DB 991 NINWRCNDNDGDSDEAGCSH 1013

RESULT 8
US-08-469-658-52
Sequence 52, Application US/08469658
Patent No. 5917018
GENERAL INFORMATION:

APPLICANT: Thøgersen, Hans Christian
APPLICANT: Holtet, Thor Las
APPLICANT: Etzerodt, Michael
TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOOLDING OF
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,658
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/192,060
FILING DATE: February 4, 1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Paul T. Clark
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06363/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 542 5070
TELEFAX: 617 542 8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 4544 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-658-52

Query Match 30.3%; Score 47.5; DB 2; Length 4544;
Best Local Similarity 39.1%; Pred. No. 4.5e-02;
Matches 9; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

QY 4 SVGWRLQNGDDMDHAICGH 23
DB 991 NINWRCNDNDGDSDEAGCSH 1013

RESULT 9
US-09-679-279-16
Sequence 16, Application US/09679279
Patent No. 6524841
GENERAL INFORMATION:
APPLICANT: McDaniel, Robert
APPLICANT: Volchegursky, Yanina

Query Match 29.38; Score 46;

Best Local Similarity 42.1%; Pred. No. 2.1; Indels 7; Gaps 0; Mismatches 4; Gaps 0;

QY 2 VRSVGWRLQIGDDMDHAI 20
Db 5 IHNIARHLAQIGDMDHNI 23

RESULT 14

US-08-706-741B-55
; Sequence 55, Application US/08706741B
; Patent No. 5955593
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, STANLEY J.
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: USA
; ZIP: 63146

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706.741B
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 965017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-706-741B-55

Query Match 29.3%; Score 46; DB 2; Length 32;
Best Local Similarity 42.1%; Pred. No. 2.6;
Matches 8; Conservative 4; Mismatches 7; Indels 7; Gaps 0;

QY 2 VRSVGWRLQIGDDMDHAI 20
Db 9 IHNIARHLAQIGDMDHNI 27

RESULT 15

US-08-924-695A-55
; Sequence 55, Application US/08924695A
; Patent No. 5998583
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, STANLEY J.
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: USA
; ZIP: 63105

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/924.695A
; FILING DATE: 09-SEP-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 971798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-924-695A-55

Query Match 29.3%; Score 46; DB 2; Length 32;
Best Local Similarity 42.1%; Pred. No. 2.6;
Matches 8; Conservative 4; Mismatches 7; Indels 7; Gaps 0;

QY 2 VRSVGWRLQIGDDMDHAI 20
Db 9 IHNIARHLAQIGDMDHNI 27

Search completed: November 10, 2004, 15:57:18
Job time : 12.9537 secs

This Page Blank (uspto)

; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 7444
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C152384_1.p
US-10-739-930-7444

Query Match 38.9%; Score 61; DB 17; Length 381;
Best Local Similarity 52.4%; Pred. No. 1.6;
Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 FVRSVGRWLQIGDDMDHAIC 21
DB 272 FAERVGWRLQKLDGNGVQAF 292

RESULT 3
US-10-424-599-247510
; Sequence 247510, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 247510
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_65530C.1.p
US-10-424-599-247510

Query Match 38.2%; Score 60; DB 15; Length 526;
Best Local Similarity 54.5%; Pred. No. 3.2;
Matches 12; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRLQIGDDMDHAICGHDVRL 27
DB 310 GWRTNRIGDDGLHAIKACLN 331

RESULT 4
US-10-425-114-49445
; Sequence 49445, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 49445
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700953440_FLI.p
US-10-425-114-49445

Query Match 38.2%; Score 60; DB 15; Length 532;
Best Local Similarity 54.5%; Pred. No. 3.2;
Matches 12; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRLQIGDDMDHAICGHDVRL 27
DB 316 GWRTNRIGDDGLHAIKACLN 337

RESULT 5
US-10-424-599-235298
; Sequence 235298, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 235298
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_54501C.1.p
US-10-424-599-235298

Query Match 37.6%; Score 59; DB 15; Length 148;
Best Local Similarity 42.9%; Pred. No. 1.2;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 FVRSVGRWLQIGDDMDHAIC 21
DB 37 FAERVGWRLQKLDGNGVQAF 57

RESULT 6
US-10-739-930-9193
; Sequence 9193, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 9193
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: GLYMA-23APR03-C5136_1.p
US-10-739-930-9193

Query Match 37.6%; Score 59; DB 17; Length 332;
Best Local Similarity 42.9%; Pred. No. 2.8;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

```
QY 1 FVRVSGWRLQNIQIGDDMDHAIC 21
Db 221 FAEKVGWIKQKRDDELLIHEIC 241

RESULT 7
US-10-424-599-246092
; Sequence 246092, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 246092
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_64251C.1.pep
US-10-424-599-246092

Query Match 35.0%; Score 55; DB 15; Length 341;
Best Local Similarity 38.1%; Pred. No. 11;
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 FVRVSGWRLQNIQIGDDMDHAIC 21
Db 221 FAEKVGWIKQKRDDELLIHEFC 241

RESULT 8
US-10-424-599-218856
; Sequence 218856, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 218856
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(479)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_39655C.1.pep
US-10-424-599-218856

Query Match 34.4%; Score 54; DB 15; Length 479;
Best Local Similarity 45.8%; Pred. No. 22;
Matches 11; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

QY 1 FVRVSGWRLQNIQIGDDMDHAICGH 24
Db 425 YADSIGRIIE--GDDVDEAPCGD 446

RESULT 9
US-10-425-115-286505
; Sequence 286505, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 286505
; LENGTH: 1551
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_24390C.1.pep
US-10-425-115-286505

Query Match 34.4%; Score 54; DB 17; Length 1551;
Best Local Similarity 42.1%; Pred. No. 73;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 VRSVGWRLQNIQIGDDMDHAI 20
Db 1104 VRAVRWRLKDLGTEVPHCV 1122

RESULT 10
US-10-424-599-153311
; Sequence 153311, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 153311
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_109464C.1.pep
US-10-424-599-153311

Query Match 33.8%; Score 53; DB 15; Length 124;
Best Local Similarity 39.1%; Pred. No. 7.8;
Matches 9; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 VRSVGWRLQNIQIGDDMDHAI 24
Db 31 IRITSWRLSVLPGSTHSQCGDD 53

RESULT 11
US-10-282-122A-50976
; Sequence 50976, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
```

```

; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50976
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-10-282-122A-50976

Query Match 33.8%; Score 53; DB 15; Length 264;
Best Local Similarity 47.6%; Pred. No. 17;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 VRSVGVRLQIGDDMDHAICG 22
DB 227 VRSVQVHVSNNVRDPDCAVCG 247

RESULT 12
; Sequence 8112, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8112
; LENGTH: 604
; TYPE: PRT
; ORGANISM: Thermobifida fusca
US-10-369-493-8112

```

```

Query Match 33.8%; Score 53; DB 14; Length 604;
Best Local Similarity 45.5%; Pred. No. 39;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 6 GWRQLQIGDDMDHAICGHDVRL 27
DB 287 GWKVTIGDDIAWVRFGDGL 308

RESULT 13
US-10-437-963-137212
; Sequence 137212, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204956
; SEQ ID NO 137212
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_38716C.1.pap
US-10-437-963-137212

Query Match 32.8%; Score 51.5; DB 16; Length 374;
Best Local Similarity 43.5%; Pred. No. 40;
Matches 10; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 5 VGVRLQIGDDMDHAICGHDVRL 27
DB 56 VEWFKSLEDGKVAH-CGHDVHV 77

RESULT 14
US-09-272-975-5
; Sequence 5, Application US/09272975
; Publication No. US2003002774A1
; GENERAL INFORMATION:
; APPLICANT: Hendrickson, Ronald C.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: TUBERCULOSIS ANTIGENS AND METHODS
; FILE REFERENCE: 210121.474
; CURRENT APPLICATION NUMBER: US/09/272,975
; CURRENT FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-272-975-5

Query Match 32.5%; Score 51; DB 10; Length 606;
Best Local Similarity 45.5%; Pred. No. 78;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRQLQIGDDMDHAICGHDVRL 27
DB 6 GWRQLQIGDDMDHAICGHDVRL 27

```

Db 287 GWRAETLGGDDIAMWRFKDGRL 308

RESULT 15

US-10-367-094-9
; Sequence 9, Application US/10367094
; Publication No. US20040170982A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001500
; CURRENT APPLICATION NUMBER: US/10/367,094
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-367-094-9

Query Match 32.2%; Score 50.5; DB 16; Length 605;
Best Local Similarity 44.4%; Pred. No. 92;
Matches 12; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

Qy 1 FVRSVGVRLQNIQDDMDHAICGHDVRL 27
Db 23 FFRGVWEELFNVGDDV-YALFGSDINL 48

Search completed: November 11, 2004, 07:41:44
Job time : 67.8584 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:52:32 ; Search time 9.56584 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-64

Perfect score: 157

Sequence: 1 FVRSVGRQLNIGDDMDHAICGHDVRLG 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 791.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	32.5	606	2 A70960	probable pckA prot
2	51	32.5	609	2 F87237	phosphoenolpyruvat
3	51	32.5	618	2 A71364	probable phosphoen
4	50	31.8	340	2 AB2584	hypothetical prote
5	50	31.8	340	2 B97366	ficB protein (AF19
6	50	31.8	599	2 G71481	probable phosphoen
7	50	31.8	600	2 E72027	phosphoenolpyruvat
8	50	31.8	600	2 A86597	phosphoenolpyruvat
9	50	31.8	608	2 JQ1462	phosphoenolpyruvat
10	49	31.2	295	2 C72692	probable potassium
11	49	31.2	786	2 G96940	hypothetical prote
12	48	30.6	108	2 G83689	hypothetical prote
13	48	30.6	262	2 C84563	hypothetical prote
14	48	30.6	492	2 T40304	spliceosome-associ
15	48	30.6	619	2 A45625	phosphoenolpyruvat
16	48	30.6	654	2 T24168	hypothetical prote
17	47.5	30.3	218	2 AB3314	hydroxyacylglutath
18	47.5	30.3	336	2 F82737	ornithine carbamoy
19	47.5	30.3	833	2 T10695	transcription fact
20	47.5	30.3	4543	1 A53102	alpha-2-macroglobu
21	47.5	30.3	4544	1 S23392	alpha-2-macroglobu
22	47.5	30.3	4545	1 S25111	probable acetyltra
23	47	29.9	205	2 C75096	hypothetical prote
24	47	29.9	374	2 AB1943	hypothetical prote
25	47	29.9	464	2 F83365	lipamide dehydrog
26	47	29.9	1249	2 C83358	probable non-ribos
27	46	28.3	199	2 S72728	probable L-a-amino
28	46	28.3	285	2 AB0106	probable lipoprote
29	46	29.3	372	1 G70817	probable ABC trans

30 46 29.3 389 2 S44367
31 46 29.3 555 2 D90354
32 46 29.3 619 2 C83168
33 46 29.3 646 2 S30180
34 46 29.3 651 2 T32875
35 46 29.3 685 2 S78040
36 46 29.3 705 2 S34968
37 46 29.3 767 2 D95101
38 46 29.3 767 2 E97969
39 45.5 29.0 363 1 W2BE44
40 45.5 29.0 384 2 B86660
41 45 28.7 129 2 A30554
42 45 28.7 159 2 B82742
43 45 28.7 178 2 E81672
44 45 28.7 286 2 T38680
45 45 28.7 296 2 A90191

ALIGNMENTS

RESULT 1

A70960

Probable pckA protein - Mycobacterium tuberculosis (strain H37Rv)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 03-Jul-2004

C;Accession: A70960

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: A70960

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-606 <COL>

A;Cross-references: UNIPROT:P96393; GB:292669; GB:AL123456; NID:G3242271; PIDN:CAB07006

A;Experimental source: strain H37Rv

C;Genetics:

A;Gene: pckA

C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match 32.5%; Score 51; DB 2; Length 606;
Best Local Similarity 45.5%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRLQNIQDDMDHAICGHDVRL 27

|||||:|||||

DB 287 GWRAETLGDDIAWMRFKDGRL 308

RESULT 2

F87237

phosphoenolpyruvate carboxykinase [imported] - Mycobacterium leprae

C;Species: Mycobacterium leprae

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: F87237

R;Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H

R.; Davies, R.M.; Devlin, K.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd

eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S

A;Title: Massive gene decay in the leprosy bacillus.

A;Reference number: A86909; MUID:21128732; PMID:11234002

A;Accession: F87237

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-609 <STO>

A;Cross-references: UNIPROT:O06084; GB:AL450380; NID:G13093848; PIDN:CAC32156.1; GSPDB

C;Genetics:

A;Gene: pckA

Qy 9 LQNI GDDMDHAICGH D VRLG 28
: | | | : | | | | |
Db 188 INKIGDEDDKRI VGH DVTNG 207

R.Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J. Nature Genet. 21, 385-389, 1999
Article: Comparative genomes of *Clamydia pneumoniae* and *C. trachomatis*.
Reference number: A72000; MCID:99206606; PMID:10192388
Accession: E72027
A:Accession: E72027
A:Status: preliminary
A:Molecule type: DNA

```
A;Residues: 1-600 <ARN>
A;Cross-references: UNIPROT:Q97755; GB:AE001666; NID:g4377164; PIDN:AAD1898
A;Experimental source: strain CWL029
C;Genetics:
A;Gene: pckA
C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match      31.8%; Score 50; DB 2; Length 600;
Best Local Similarity 45.5%; Pred. No. 18;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy      6  GWRLLQIGDDMDHAICGHVDVRL 27
Db      281  GWKIEICIGDDIAWIRPGRDGRL 302

RESULT 8
A86597
phosphoenolpyruvate carboxykinase [imported] - Chlamydothila pneumoniae (strain J138)
C;Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A86597
R;Shirai, M.; Hirakawa, H.; Kinoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise, N.
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: A86597
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-600 <STO>
A;Cross-references: UNIPROT:Q97755; GB:BA000008; NID:g9979225; PIDN:BAA99059.1; GSPDB:GN
A;Experimental source: strain J138
C;Genetics:
A;Gene: pckA
C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)

Query Match      31.8%; Score 50; DB 2; Length 600;
Best Local Similarity 45.5%; Pred. No. 18;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy      6  GWRLLQIGDDMDHAICGHVDVRL 27
Db      281  GWKIEICIGDDIAWIRPGRDGRL 302

RESULT 9
JQ1462
phosphoenolpyruvate carboxykinase (GTP) (BC 4.1.1.32) - rumen fungus (Neocallimastix fr
C;Species: Neocallimastix frontalis
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C;Accession: JQ1462
R;Reymond, P.; Geourjon, C.; Roux, B.; Durand, R.; Fevre, M.
Gene 110, 57-63, 1992
A;Title: Sequence of the phosphoenolpyruvate carboxykinase-encoding cDNA from the rumen
A;Reference number: JQ1462; MUID:92184115; PMID:1339359
A;Accession: JQ1462
A;Molecule type: mRNA
A;Residues: 1-608 <REV>
A;Cross-references: UNIPROT:P22130; GB:M59372; NID:g168741; PIDN:AAA33553.1; PID:g168742
C;Comment: This enzyme is a gluconeogenic enzyme which catalyzes the reversible conversi
C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)
C;Keywords: carbon-carbon lyase; carboxy-lyase; gluconeogenesis

Query Match      31.8%; Score 50; DB 2; Length 608;
Best Local Similarity 40.9%; Pred. No. 18;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy      6  GWRLLQIGDDMDHAICGHVDVRL 27
Db      283  GWKIECVGDDIAWLKIGKDGRL 304

RESULT 10
```

```
C72692
Probable potassium channel APE0955 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: C72692
R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: C72692
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-295 <KAW>
A;Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA79939.1; PID:1043725; PID:g51
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0955

Query Match      31.2%; Score 49; DB 2; Length 295;
Best Local Similarity 45.8%; Pred. No. 12;
Matches 11; Conservative 5; Mismatches 4; Indels 4; Gaps 1;

Qy      5  VGRLLQIGDDMDHAICGHVDVRLG 28
Db      24  LGRVRNIGDVMEHPL----VELG 43

RESULT 11
G96940
hypothetical protein CAC0333 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: G96940
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G96940
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-786 <KUR>
A;Cross-references: UNIPROT:Q97M65; GB:AE001437; PIDN:AAK78314.1; PID:g15023179; GSPDB:
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0333

Query Match      31.2%; Score 49; DB 2; Length 786;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      9  LQNIQIGDDMDHAICGHVDVRL 26
Db      654  IQKIGMDMDHASNQYFLR 671

RESULT 12
G83689
hypothetical protein BH0319 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: G83689
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: G83689
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-108 <STO>
A;Cross-references: UNIPROT:Q9KPF7; GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BA804
A;Experimental source: strain C-125
```

C;Genetics:

A;Gene: BH0319

Query Match

30.6%; Score 48; DB 2; Length 108;

Best Local Similarity

44.0%; Pred. No. 5.8;

Matches 11; Conservative

3; Mismatches 7; Indels 4; Gaps 1;

QY 6 GWRLQNIQ---

DDMDHAICGHDR 26

DB 5 GWRLNSSSPAVNDLEHHIKGHFR 29

RESULT 13

C84563

hypothetical protein At2g18350 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C;Accession: C84563

R;Bin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: C84563

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-262 <L>STO>

A;Cross-references: UNIPROT:Q9ZPW7; GB:AE002093; NID:g4309732; PIDN:AA15502.1; GSPDB:GN

C;Genetics:

A;Gene: At2g18350

A;Map position: 2

Query Match

30.6%; Score 48; DB 2; Length 262;

Best Local Similarity

33.3%; Pred. No. 15;

Matches 7; Conservative

4; Mismatches 10; Indels 0; Gaps 0;

QY 1 FVRSVGWRLQNIQIGDDMDHAIC 21

DB 216 FAEKIGWRTKLEDDDEVNRFC 236

RESULT 14

T40304

spliceosome-associated protein - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T40304

R;Byrne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Moestl, D.; Duesterh

submitted to the EMBL Data Library, May 1998

A;Reference number: Z21919

A;Accession: T40304

A;Status: Preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-492 <L>LYN>

A;Cross-references: UNIPROT:Q59706; EMBL:AL023589; PIDN:CAA19057.1; GSPDB:GN00067; SPDB:

A;Experimental source: strain 972h-; cosmid c36

C;Genetics:

A;Gene: SPDB:SPBC36.09

A;Map position: 2

A;Introns: 67/2; 449/1

Query Match

30.6%; Score 48; DB 2; Length 492;

Best Local Similarity

36.4%; Pred. No. 29;

Matches 8; Conservative

5; Mismatches 9; Indels 0; Gaps 0;

QY 7 WRLQNIQIGDDMDHAICGHDRVLG 28

DB 388 WKLHGLGKEFPCEICGNVYMG 409

RESULT 15

A45625

phosphoenolpyruvate carboxykinase (GTP) (EC 4.1.1.32) - nematode (Haemonchus contortus)

N;Alternate names: phosphoenolpyruvate carboxylase

C;Species: Haemonchus contortus

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004

C;Accession: A45625

R;Klein, R.D.; Winterrowd, C.A.; Hatzenbuehler, N.T.; Shea, M.H.; Favreau, M.A.; Nulff, S.

Mol. Biochem. Parasitol. 50, 285-294, 1992

A;Title: Cloning of a cDNA encoding phosphoenolpyruvate carboxykinase from Haemonchus c

A;Reference number: A45625; MUID:92158009; PMID:1741016

A;Accession: A45625

A;Status: Preliminary

A;Molecule type: mRNA

A;Residues: 1-619 <KLE>

A;Cross-references: UNIPROT:P29190; GB:M76494; NID:gl59182; PIDN:AAA29180.1; PID:gl5918

A;Note: sequence extracted from NCBI backbone (NCBIN:83527, NCBIP:83529)

C;Superfamily: phosphoenolpyruvate carboxykinase (GTP)

C;Keywords: carbon-carbon lyase; carboxy-lyase; gluconeogenesis

Query Match

30.6%; Score 48; DB 2; Length 619;

Best Local Similarity

40.9%; Pred. No. 37;

Matches 9; Conservative

5; Mismatches 8; Indels 0; Gaps 0;

QY 6 GWRLQNIQIGDDMDHAICGHDRVL 27

DB 297 GWKVCVGGDDIAWMKFGEDGRL 318

Search completed: November 10, 2004, 15:55:02

Job time : 10.5658 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 46.9324 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-64

Perfect score: 157

Sequence: 1 FVRSVGWRLQIGDDMDHAICGHVRLG 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	38.9	230	2 Q70K64	Q70K64 gordonia we
2	61	38.9	230	2 CAE09117	CAE09117 gordonia
3	54	34.4	1067	2 Q874G3	Q874G3 saccharomyc
4	53	33.8	264	2 Q7VS40	Q7VS40 bordetella
5	53	33.8	264	2 Q7W1Q3	Q7W1Q3 bordetella
6	53	33.8	264	2 Q7WQW9	Q7WQW9 bordetella
7	52	33.1	605	1 PFCK MYCSM	Q9AGJ6 mycobacteri
8	52	33.1	609	2 Q6F8P2	Q6F8P2 acinetobact
9	52	33.1	2507	2 Q81HK8	Q81HK8 dictyosteli
10	52	33.1	2929	2 Q86JG5	Q86JG5 dictyosteli
11	51.5	32.8	417	2 Q851L5	Q851L5 oryza sativ
12	51.5	32.8	417	2 AAR8S567	AAR8S567 oryza sat
13	51	32.5	606	1 PFCK MYCTU	P96393 mycobacteri
14	51	32.5	609	1 PFCK MYCLE	O06084 mycobacteri
15	51	32.5	609	2 Q73TS2	Q73TS2 mycobacteri
16	51	32.5	609	2 AAS06196	AAS06196 mycobacte
17	51	32.5	618	1 PFCK TREPA	Q33159 treponena p
18	51	32.5	814	2 Q98SW5	Q98SW5 xenopus lae
19	50.5	32.2	328	2 Q80PF9	Q80PF9 suid herpes
20	50.5	32.2	601	2 Q99M67	Q99M67 mus musculu
21	50.5	32.2	602	2 Q7TMW0	Q7TMW0 mus musculu
22	50.5	32.2	821	1 SYI_THTH	P56690 thermus the
23	50.5	32.2	1067	2 Q72JF6	Q72JF6 thermus the
24	50.5	32.2	1067	2 AAS81050	AAS81050 thermus t
25	50	31.8	202	2 Q98P24	Q98P24 rhizobium l
26	50	31.8	253	2 Q83XW9	Q83XW9 streptomyc
27	50	31.8	340	2 Q7D298	Q7D298 agrobacteri
28	50	31.8	340	2 Q8UJ75	Q8UJ75 agrobacteri
29	50	31.8	478	2 Q88LK0	Q88LK0 pseudomonas
30	50	31.8	599	1 PFCK CHLTR	Q84716 chlamydia t
31	50	31.8	600	1 PFCK_CHLPN	Q92755 chlamydia p

RESULT 1

Q70K64

ID Q70K64 PRELIMINARY; PRT; 230 AA.

AC Q70K64;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.

OS Gordonia westfalica.

OG Plasmid pKb1.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Gordoniaceae; Gordonia.

OX NCBI_TaxID:158898;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-type strain: DSM44215;

RA Broecker D., Arenscoetter M., Legatzki A., Nies D.H., Steinbuechel A.;

RT "Characterization of the 101.016 kbp megaplasmid pKb1 isolated from

RT the rubber degrading bacterium Gordonia westfalica Kb1.";

RL J. Bacteriol. 186:212-225(2004).

DR EMBL; AJ576039; CAE09117.1; -

KW Hypothetical protein; Plasmid.

SQ SEQUENCE 230 AA; 25736 MW; 923260575B53D0F2 CRC64;

Query Match 38.9%; Score 61; DB 2; Length 230;

Best Local Similarity 45.8%; Pred. No. 0.96;

Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 FVRSVGWRLQIGDDMDHAICGH 24

Db 130 FCRGCGWHSEVVGDDTDAALLGLD 153

RESULT 2

CAE09117

ID CAE09117 PRELIMINARY; PRT; 230 AA.

AC CAE09117;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.

OS Gordonia westfalica.

OG Plasmid pKb1.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Gordoniaceae; Gordonia.

OX NCBI_TaxID:158898;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-type strain: DSM44215;

RA Broecker D., Arenscoetter M., Legatzki A., Nies D.H., Steinbuechel A.;

RT "Characterization of the 101.016 kbp megaplasmid pKb1 isolated from

RT the rubber degrading bacterium Gordonia westfalica Kb1.";

RL J. Bacteriol. 186:212-225(2004).

RN [2]

Q82171 streptomyc
P22130 neocallimas
Q93115 streptomyc
Q7N482 photorhabdu
Q9Y9F8 aeropyrum p
Q985P8 rhizobium l
Q6MBF4 parachlamy
Caf24095 parachlam
Q97M65 clostridium
Q6D580 erwinia car
Q8C9U6 mus musculu
Q87F59 bradyrhizob
Q7XTV7 oryza sativ
Q9KFZ7 bacillus ha

RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtmann M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norbercrak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmons M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.",
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640423; CAE30039.1; -.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR007901; MoeZ_MoeB.
DR InterPro; IPR002005; NAD_BS.
DR InterPro; IPR000594; ThiF domain.
DR Pfam; PF05237; MoeZ_MoeB; 1.
DR Pfam; PF00899; ThiF; 1.
DR Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 264 AA; 28094 MW; 1FC05B44B6A9BDC5 CRC64;
Query Match 33.8%; Score 53; DB 2; Length 264;
Best Local Similarity 47.6%; Pred. No. 17;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 2 VRSVGWRLQNGDMDHAIICG 22
Db 227 VRSQMWHSVNVRPDPDCAVCG 247
RESULT 6
ID Q7WQM9 PRELIMINARY; PRT; 264 AA.
AC Q7WQM9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Adenylyltransferase (EC 2.7.7.-)
GN Name=thiF; OrderedLocNames=BB0301;
OC Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
FN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngi1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtmann M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norbercrak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmons M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.",
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640437; CAE30799.1; -.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR007901; MoeZ_MoeB.
DR InterPro; IPR002005; NAD_BS.
DR Pfam; PF05237; MoeZ_MoeB; 1.
DR Pfam; PF00899; ThiF; 1.
DR Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 264 AA; 28124 MW; 60B8EDB37F002E4 CRC64;
Query Match 33.8%; Score 53; DB 2; Length 264;

Best Local Similarity 47.6%; Pred. No. 17;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 2 VRSVGWRLQNGDMDHAIICG 22
Db 227 VRSQMWHSVNVRPDPDCAVCG 247
RESULT 7
ID PPCK_MYCSM STANDARD; PRT; 605 AA.
AC Q9AGJ6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE phosphoenolpyruvate carboxykinase [GTP] (EC 4.1.1.32) (PEP
DE carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).
GN Name=ppcK; Synonyms=ppk;
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
FN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND MASS SPECTROMETRY.
RX MEDLINE=21238287; PubMed=11278451; DOI=10.1074/jbc.M008960200;
RA Mukhopadhyay B., Concar E.M., Wolfe R.S.;
RT "A GTP-dependent vertebrate-type phosphoenolpyruvate carboxykinase
RT from Mycobacterium smegmatis.",
RL J. Biol. Chem. 276:16137-16145(2001).
CC -|- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2).
CC -|- ENZYME REGULATION: Inhibited by oxalate and by alpha-
CC ketoglutarate.
CC -|- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -|- SUBUNIT: Monomer.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- MASS SPECTROMETRY: MW=71209; METHOD=MALDI; RANGE=1-605;
CC NOTE=Ref.1.
CC -|- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [GTP]
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF332191; AAK28534.1; -.
CC HSSP; P35558; 1KHB.
CC HAMAP; MF 00452; 1.
CC InterPro; IPR008210; PEPCK_N.
CC InterPro; IPR008289; PEP_Carboxykin.
CC Pfam; PF00821; PEPCK; 1.
CC ProDom; PD004738; PEPCK_N; 1.
CC PROSITE; PS00505; PEPCK_GTP; 1.
CC Decarboxylase; Gluconeogenesis; GTP-binding; Lyase.
FT NP_BIND 222 229 GTP (Potential).
FT ACT_SITE 273 273 By similarity.
SQ SEQUENCE 605 AA; 66943 MW; E07A46B4F35EDC0 CRC64;
Query Match 33.1%; Score 52; DB 1; Length 605;
Best Local Similarity 45.5%; Pred. No. 58;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 6 GWRLQNGDMDHAIICGHVRL 27
Db 287 GWRAETVGGDDIAWNRFGKGRLL 308
RESULT 8
Q6F8P2

```

ID O6F8P2 PRELIMINARY; PRT; 609 AA.
AC O6F8P2;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Phosphoenolpyruvate carboxykinase (EC 4.1.1.32).
GN Name: pckG; Ordered locus names: ACIAD2842;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP SEQUENCE FROM N.A.
RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Ornston L.N., Weissbach J., Marliere P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium."
RL Nucleic Acids Res. 0:0-0(2004).
DR EMBL; CR543861; CAG69573.1; -.
DR InterPro; IPR008210; PEPCK_N.
DR InterPro; IPR008209; PEP_carboxykin.
DR Pfam; PF00821; PEPCK; 1_carboxykin.
DR ProDom; PD004738; PEPCK_N; 1.
DR PROSITE; PS00505; PEPCK_GTP; 1.
KW Complete proteome; Kinase; Lyase; Pyruvate.
SQ SEQUENCE 609 AA; 67306 MW; 5301EBFB66E3231E CRC64;

Query Match 33.1%; Score 52; DB 2; Length 609;
Best Local Similarity 40.9%; Pred. No. 58;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 6 GWRQNIGDDMDHAIICGHVRL 27
Db 285 GWKIETVGDIAMIKPGDGRL 306
|||:|||||
|||:|||||

RESULT 9
Q8IHK8 PRELIMINARY; PRT; 2507 AA.
AC Q8IHK8;
DT 01-WAR-2003 (T-EMBLrel. 23, Created)
DT 01-WAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-WAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Lvsd (Fragment).
GN Name: lvsd;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris E., Wang N., Wu W.L., Weatherford A., De Loranne A.,
RA Cardelli J.;
RT "Dictyostelium LvsB mutants model the lysosomal defects associated
RT with Chediak-Higashi syndrome."
RL Mol. Biol. Cell 13:656-669(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE-21843693; PubMed=11854420;
RA Harris E., Wang N., Wu W.L., Weatherford A., De Loranne A.,
RA Cardelli J.;
RT "Dictyostelium LvsB mutants model the lysosomal defects associated
RT with Chediak-Higashi syndrome."
RL Mol. Biol. Cell 13:656-669(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE-22198669; PubMed=12210762;
RA Wang N., Wu W.L., De Loranne A.;
RT "BEACH family of proteins: phylogenetic and functional analysis of six
RT Dictyostelium BEACH proteins."
RL J. Cell. Biochem. 86:561-570(2002).
CC -1- SIMILARITY: Contains 2 WD repeats.
DR EMBL; AV159040; AAN38987.1; -.
DR DictyBase; DB0185108; lvsd.
DR InterPro; IPR000403; Beige_BEACH.
DR InterPro; IPR000508; Peptidase_S26.
DR InterPro; IPR011680; WD40.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF02138; Beach; 1.
DR Pfam; PF00400; WD40; 2.

```

```

DR SMART; SM00320; WD40; 2.
DR PROSITE; PS0197; BEACH; 1.
DR PROSITE; PS00501; SPASE_I_1; UNKNOWN_1.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT NON_TER 1
SQ SEQUENCE 2507 AA; 283343 MW; 3821DD8B23CE476 CRC64;

Query Match 33.1%; Score 52; DB 2; Length 2507;
Best Local Similarity 42.9%; Pred. No. 25e+02;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 7 WRLQNGDDMDHAIICGHVRL 27
Db 1016 WGLPMISEDMDHIISKNSIRI 1036
|||:|||||
|||:|||||

RESULT 10
Q86JGS PRELIMINARY; PRT; 2929 AA.
AC Q86JGS;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Similar to Dictyostelium discoideum (Slime mold). Lvsd.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RC MEDLINE-22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggall B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum."
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RC Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 2 WD repeats.
DR EMBL; AC116982; AAC51610.1; -.
DR HSP; O8NFP9; 1M1.
DR InterPro; IPR000409; Beige_BEACH.
DR InterPro; IPR005308; Peptidase_S26.
DR InterPro; IPR011680; WD40.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF02138; Beach; 1.
DR Pfam; PF00400; WD40; 2.
DR SMART; SM00320; WD40; 2.
DR PROSITE; PS0197; BEACH; 1.
DR PROSITE; PS00501; SPASE_I_1; UNKNOWN_1.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 2929 AA; 331654 MW; 0E822ED10B77A19D CRC64;

Query Match 33.1%; Score 52; DB 2; Length 2929;
Best Local Similarity 42.9%; Pred. No. 2.9e+02;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 7 WRLQNGDDMDHAIICGHVRL 27
Db 1438 WGLPMISEDMDHIISKNSIRI 1458
|||:|||||
|||:|||||

RESULT 11
Q851L5 PRELIMINARY; PRT; 417 AA.
AC Q851L5;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)

```


05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DT Putative amidohydrolase.
DE Name=OSUNBA0042109.29; Synonyms=OSUNBA0096106.2;
GN Oryza sativa (japonica cultivar-group).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidaceae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
OX [1]
RN
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
[3]
RN
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[4]
RN
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL; AC104487; AAC041148.1; -
DR ENBL; AC092557; AAR88567.1; -
DR Gramene; Q8SIL5; -
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR GO; GO:0008237; F:metalloproteinase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002933; Peptidase M20.
DR InterPro; IPR010168; Pept_M20D_amidh.
DR Pfam; PF01546; Peptidase_N20; I.
DR TIGRFAMs; TIGR01891; amidohydrolases; 1.
KW Hydrolase.
KW
SQ SEQUENCE 417 AA; 49990 MW; 7F1326A1FB883B54 CRC64;

Query Match 32.8%; Score 51.5; DB 2; Length 417;
Best Local Similarity 43.5%; Pred. No. 47;
Matches 10; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY 5 VGWRLQNIGDDMDHAICGHVRL 27
| | : : : | | | | :
Db 99 VEWEFKSLDGKXMA-CGHVDHV 120

RESULT 12
AAR88567 ID AAR88567 PRELIMINARY; PRT; 417 AA.
AC AAR88567;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative amidohydrolase.
GN OSUNBA0096106.2.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidaceae; Oryzaceae; Oryza; Oryza sativa.
OX NCBI_TaxID=39947;
OX [1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;

```

RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garner T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2).
CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [GTP]
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; BX842572; CAB07006.1; -.
CC EMBL; AE006931; AAK44442.1; -.
CC EMBL; BX248334; CAD93081.1; -.
CC PIR; A70960; A70960.
CC HSSP; P35558; IKBH.
CC TIGR; MT0221; -.
CC
CC TubercuList; RV0211; -.
CC HAMAP; MF_00452; -.
CC InterPro; IPR008210; PEPCK_N.
CC InterPro; IPR008209; PEP_carboxykin.
CC Pfam; PF00821; PEPCK_1.
CC ProDom; PD004738; PEPCK_N; 1.
CC PROSITE; PS00505; PEPCK_GTP; 1.
CC PROSITE; PS00505; PEPCK_N; 1.
CC Complete proteome; Decarboxylase; Gluconeogenesis; GTP-binding; Lyase.
FT NP BIND 222 229 GTP (Potential).
FT ACT SITE 273 273 By similarity.
SQ SEQUENCE 606 AA; 67253 MW; AEE29412E6BCCAE3 CRC64;

Query Match 32.5%; Score 51; DB 1; Length 606;
Best Local Similarity 45.5%; Pred. No. 82;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRLQNGDDMDHAIICGHVDRL 27
Db 287 GWRAETLGDDIAWMRFKDGRL 308

RESULT 14
PPCK_MYCLE
ID PPCK_MYCLE STANDARD; PRT; 609 AA.
AC Q06084;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Phosphoenolpyruvate carboxykinase [GTP] (EC 4.1.1.32) (PEP
DE carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).
GN Name=pckG; Synonyms=pckA; OrderedLocNames=ML2624;
GN ORFNames=MLC622.21;
OS Mycobacterium leprae
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,
RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,

```

```

RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2).
CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [GTP]
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z95398; CAB08805.1; -.
CC EMBL; AL583926; CAC32156.1; -.
CC PIR; F87237; F87237.
CC HSSP; P35558; IKBH.
CC Leproma; ML2624; -.
CC HAMAP; MF_00452; -.
CC InterPro; IPR008210; PEPCK_N.
CC InterPro; IPR008209; PEP_carboxykin.
CC Pfam; PF00821; PEPCK_1.
CC ProDom; PD004738; PEPCK_N; 1.
CC PROSITE; PS00505; PEPCK_GTP; 1.
CC PROSITE; PS00505; PEPCK_N; 1.
CC Complete proteome; Decarboxylase; Gluconeogenesis; GTP-binding; Lyase.
FT NP BIND 222 229 GTP (Potential).
FT ACT SITE 273 273 By similarity.
SQ SEQUENCE 609 AA; 67731 MW; 8916B6733622590B CRC64;

Query Match 32.5%; Score 51; DB 1; Length 609;
Best Local Similarity 45.5%; Pred. No. 82;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 6 GWRLQNGDDMDHAIICGHVDRL 27
Db 287 GWRAETLGDDIAWMRFKDGRL 308

RESULT 15
Q73TS2
ID Q73TS2 PRELIMINARY; PRT; 609 AA.
AC Q73TS2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE PckA.
GN Name=pckA; OrderedLocNames=MAP3646;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K10;
RA Li L., Barnantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (Sep-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AR017240; AAS06196.1; -.
DR InterPro; IPR008210; PEPCK_N.
DR InterPro; IPR008209; PEP_carboxykin.
DR Pfam; PF00821; PEPCK_1.
DR ProDom; PD004738; PEPCK_N; 1.
DR PROSITE; PS00505; PEPCK_GTP; 1.

```

KW Complete proteome.
SQ SEQUENCE 609 AA; 67659 MW; 8ABF04BFFF2D1847 CRC64;
Query Match 32.5%; Score 51; DB 2; Length 609;
Best Local Similarity 45.5%; Pred. No. 82;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 6 GWRLLQNIIGDDMDHAICGHVRL 27
Db ||| : |||
287 GWRATLGGDDIAWVRFGKGRLL 308

Search completed: November 10, 2004, 15:53:22
Job time : 49.0093 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 6.01424 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-65
Perfect score: 69
Sequence: 1 SGURKPTCGSSQR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41.5	60.1	501	4	US-09-252-991A-23033
2	40	58.0	494	1	US-07-906-930E-33
3	38	55.1	65	4	US-09-270-767-44530
4	38	55.1	153	4	US-09-252-991A-26199
5	38	55.1	163	4	US-09-252-991A-24968
6	38	55.1	469	4	US-09-252-991A-25490
7	38	55.1	619	4	US-09-252-991A-21585
8	37	53.6	88	4	US-09-513-999C-5920
9	37	53.6	136	4	US-09-252-991A-19336
10	37	53.6	269	4	US-09-252-991A-29509
11	37	53.6	362	4	US-09-489-039A-8597
12	37	53.6	459	4	US-09-252-991A-20121
13	37	53.6	1299	5	PCT-US95-08354A-2
14	36.5	52.9	61	2	US-08-744-670-6
15	36.5	52.9	61	2	US-09-149-933-6
16	36	52.2	50	4	US-09-621-976-5086
17	36	52.2	74	2	US-08-463-360-42
18	36	52.2	74	2	US-08-486-397-42
19	36	52.2	74	2	US-08-486-399-42
20	36	52.2	74	2	US-08-461-965-42
21	36	52.2	74	2	US-08-634-641-42
22	36	52.2	74	3	US-09-249-471-42
23	36	52.2	74	3	US-09-249-472-42
24	36	52.2	74	3	US-09-249-451-42
25	36	52.2	74	3	US-08-809-455-42
26	36	52.2	74	3	US-09-249-461-42
27	36	52.2	74	3	US-09-249-448-42

Sequence 42, Appl
Sequence 31721, A
Sequence 3298, A
Sequence 2284, A
Sequence 29948, A
Sequence 24012, A
Sequence 24396, A
Sequence 24667, A
Patent No. 5183745
Sequence 7, Appl
Sequence 19069, A
Patent No. 5183745
Sequence 4, Appl
Sequence 2, Appl
Patent No. 5183745
Sequence 36, Appl
Sequence 37, Appl
Sequence 22, Appl

74 4 US-09-249-473-42
138 4 US-09-252-991A-31721
222 4 US-09-252-991A-32998
424 4 US-09-252-991A-22384
427 4 US-09-252-991A-29948
441 4 US-09-252-991A-24012
441 4 US-09-252-991A-24396
461 4 US-09-252-991A-24667
525 6 5183745-4
585 4 US-09-173-300-7
692 4 US-09-252-991A-19069
1489 6 5183745-2
1705 3 US-08-669-785-4
1706 3 US-08-669-785-2
1794 6 5183745-6
8 3 US-09-020-065A-36
10 3 US-09-020-065A-37
83 2 US-08-245-511-22

ALIGNMENTS

RESULT 1
US-09-252-991A-23033
; Sequence 23033, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23033
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23033

Query Match 60.1%; Score 41.5; DB 4; Length 501;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 2 GLRKP-TCGSSQ 12
Db 69 GFRKPTCGSSR 80

RESULT 2
US-07-906-930E-33
; Sequence 33, Application US/07906930E
; Patent No. 5534631
; GENERAL INFORMATION:
; APPLICANT: Gaynor, Richard B.
; APPLICANT: Nirula, Ajay
; APPLICANT: Li, Ching
; TITLE OF INVENTION: DNA ENCODING THE INTERLEUKIN BINDING
; TITLE OF INVENTION: FACTOR (ILF)
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,930E
; FILING DATE: 30-JUN-1992
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sertich, Gary J.
; REGISTRATION NUMBER: 34,430
; REFERENCE/DOCKET NUMBER: UTSD:262/SER
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-418-3000
; TELEFAX: 512-474-7577
; TELEX: NOT APPLICABLE
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-906-930E-33

Query Match 58.0%; Score 40; DB 1; Length 494;
Best Local Similarity 70.0%; Pred. No. 62;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GLRKPTCGSS 11
Db 479 GLRRPCCASS 488

RESULT 3
US-09-270-767-44530
; Sequence 44530, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7328-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 44530
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-44530

Query Match 55.1%; Score 38; DB 4; Length 65;
Best Local Similarity 61.5%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSOR 13
Db 52 SGLRKGEGRGGOR 64

RESULT 4
US-09-252-991A-26199
; Sequence 26199, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25490
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25490

Query Match 55.1%; Score 38; DB 4; Length 469;

```

```

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26199
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26199

Query Match 55.1%; Score 38; DB 4; Length 153;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 RKPTCGSSOR 13
Db 87 RRPNCGSRR 96

RESULT 5
US-09-252-991A-24968
; Sequence 24968, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24968
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24968

Query Match 55.1%; Score 38; DB 4; Length 163;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 RKPTCGS 10
Db 66 RRPTCGS 72

RESULT 6
US-09-252-991A-25490
; Sequence 25490, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25490
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25490

Query Match 55.1%; Score 38; DB 4; Length 469;

```

Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 RKPTCGSSQR 13
Db 113 RKPTCSATSR 122

RESULT 7
US-09-252-991A-21585
; Sequence 21585, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21585
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21585

Query Match 55.1%; Score 38; DB 4; Length 619;
Best Local Similarity 77.8%; Pred. No. 1.7e-02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GLRKPTCS 10
Db 412 GLRPTSGS 420

RESULT 8
US-09-513-999C-5920
; Sequence 5920, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5920
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 31
; OTHER INFORMATION: Xaa=Pro or Thr
US-09-513-999C-5920

Query Match 53.6%; Score 37; DB 4; Length 88;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LRKPTCG 9
Db 80 LRKPTCG 86

RESULT 9
US-09-252-991A-19336
; Sequence 19336, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19336
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19336

Query Match 53.6%; Score 37; DB 4; Length 136;
Best Local Similarity 53.8%; Pred. No. 57;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSQR 13
Db 67 AGLRPPRCSSRAR 79

RESULT 10
US-09-252-991A-29509
; Sequence 29509, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29509
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29509

Query Match 53.6%; Score 37; DB 4; Length 269;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSQ 12
Db 204 AGEKKPLCGGQ 215

RESULT 11
US-09-489-039A-8597
; Sequence 8597, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 8597
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 ; US-09-489-039A-8597

Query Match 53.6%; Score 37; DB 4; Length 362;
 Best Local Similarity 55.6%; Pred. No. 1.5e+02;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 5 KPTCGSSQR 13
 :||||:|
 Db 168 EPTCGATER 176

RESULT 12

; US-09-252-991A-20121
 ; Sequence 20121, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196-136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 20121
 ; LENGTH: 459
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (46)
 ; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
 ; US-09-252-991A-20121

Query Match 53.6%; Score 37; DB 4; Length 459;
 Best Local Similarity 46.2%; Pred. No. 1.9e+02;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 SGLRPTCGSSQR 13
 :|||:|:|
 Db 8 AGLRRPPAGAERR 20

RESULT 13

; PCT-US95-08354A-2
 ; Sequence 2, Application PC/TUS9508354A
 ; GENERAL INFORMATION:
 ; APPLICANT: Temple University - Of The
 ; APPLICANT: Commonwealth System of Higher Education
 ; TITLE OF INVENTION: JAK3 PROTEIN TYROSINE
 ; TITLE OF INVENTION: KINASE AND DNA ENCODING THE SAME
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Seidel, Gonda, Lavorigna
 ; ADDRESSEE: & Monaco, P.C.
 ; STREET: Suite 1800, Two Penn Center
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19102
 ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/08354A
 ; FILING DATE:

; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/272,368
 ; FILING DATE: 8 July 1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Monaco, Daniel A.
 ; REGISTRATION NUMBER: 30,480
 ; REFERENCE/DOCKET NUMBER: 6056-203 PC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-8383
 ; TELEFAX: (215) 568-5549
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1299 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single stranded
 ; TOPOLOGY: linear
 ; PCT-US95-08354A-2

Query Match 53.6%; Score 37; DB 5; Length 1299;
 Best Local Similarity 50.0%; Pred. No. 5.1e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 LRKPTCGSSQ 12
 :|||:|:|
 Db 489 IKQPTCGSGR 498

RESULT 14

; US-08-744-670-6
 ; Sequence 6, Application US/08744670
 ; Patent No. 5858710
 ; GENERAL INFORMATION:
 ; APPLICANT: Bardman, Olga
 ; APPLICANT: Goli, Surya K.
 ; APPLICANT: Murry, Lynn E.
 ; TITLE OF INVENTION: TUMOR-ASSOCIATED KAZAL INHIBITOR
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/744,670
 ; FILING DATE: Filed Herewith
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0155 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 61 amino acids

Job time : 7.01424 secs

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: OWL
CLONE: IPST ANGAN
US-08-744-670-6

Query Match 52.9%; Score 36.5; DB 2; Length 61;
Best Local Similarity 80.0%; Pred. NO. 32;
Matches 8; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 SGL-RKPTCG 9
||| |||:
Db 4 SGLYKPCG 13

RESULT 15

US-09-149-933-6
; Sequence 6, Application US/09149933
; Patent No. 5958699
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: TUMOR-ASSOCIATED KAZAL INHIBITOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/149,933
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0155 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: OWL
; CLONE: IPST ANGAN
US-09-149-933-6

Query Match 52.9%; Score 36.5; DB 2; Length 61;
Best Local Similarity 80.0%; Pred. NO. 32;
Matches 8; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 SGL-RKPTCG 9
||| |||:
Db 4 SGLYKPCG 13

Search completed: November 10, 2004, 15:57:19

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 31.4128 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-65

Perfect score: 69

Sequence: 1 SGLRKPTCGSSQR 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69	100.0	13	14	US-10-092-750-65
2	49	71.0	95	16	Sequence 85, Appl
3	43	62.3	53	17	Sequence 138965,
4	41	59.4	323	16	Sequence 100236,
5	41	59.4	639	16	Sequence 109125,
6	40	58.0	90	17	Sequence 138962,
7	40	58.0	103	17	Sequence 312228,
8	40	58.0	137	16	Sequence 338616,
9	40	58.0	137	15	Sequence 161045,
10	40	58.0	245	15	Sequence 45264, A
11	40	58.0	291	17	Sequence 198860,
12	40	58.0	383	17	Sequence 256799,
13	40	58.0	734	16	Sequence 122034,
14	40	58.0	740	14	Sequence 2840, Ap

14	40	58.0	1011	16	US-10-408-765A-1557	Sequence 1557, Ap
15	40	58.0	1225	14	US-10-177-233-332	Sequence 332, App
16	40	58.0	1239	13	US-10-007-805-577	Sequence 577, App
17	40	58.0	1239	14	US-10-076-622-577	Sequence 577, App
18	40	58.0	1239	14	US-10-124-805-577	Sequence 577, App
19	40	58.0	19662	15	US-10-084-846A-6	Sequence 6, Appl
20	39	56.5	88	16	US-10-437-963-172364	Sequence 172364,
21	39	56.5	92	17	US-10-425-115-198430	Sequence 198430,
22	39	56.5	93	17	US-10-425-115-346930	Sequence 346930,
23	39	56.5	98	17	US-10-425-115-198429	Sequence 198429,
24	39	56.5	108	17	US-10-425-115-198428	Sequence 198428,
25	39	56.5	136	15	US-10-425-114-60501	Sequence 60501, A
26	39	56.5	153	17	US-10-425-115-198431	Sequence 198431,
27	39	56.5	297	15	US-10-425-114-67918	Sequence 67918, A
28	38.5	55.8	153	14	US-10-094-749-2445	Sequence 2445, Ap
29	38	55.1	58	17	US-10-425-115-209990	Sequence 209990,
30	38	55.1	58	17	US-10-425-115-219968	Sequence 219968,
31	38	55.1	71	17	US-10-425-115-350852	Sequence 350852,
32	38	55.1	96	16	US-10-437-963-198914	Sequence 198914,
33	38	55.1	120	9	US-09-764-847-714	Sequence 714, App
34	38	55.1	120	14	US-10-092-154-714	Sequence 714, App
35	38	55.1	123	15	US-10-424-599-234209	Sequence 234209,
36	38	55.1	149	17	US-10-425-115-264718	Sequence 264718,
37	38	55.1	191	16	US-10-437-963-183027	Sequence 183027,
38	38	55.1	195	17	US-10-425-115-310687	Sequence 310687,
39	38	55.1	276	14	US-10-302-267-164	Sequence 164, App
40	38	55.1	276	15	US-10-412-699B-664	Sequence 664, App
41	38	55.1	323	14	US-10-205-823-62	Sequence 62, Appl
42	38	55.1	323	14	US-10-177-293-53	Sequence 53, Appl
43	38	55.1	596	16	US-10-437-963-120842	Sequence 120842,
44	38	55.1	767	14	US-10-369-493-3550	Sequence 3550, Ap
45	38	55.1	807	15	US-10-108-260A-4086	Sequence 4086, Ap

ALIGNMENTS

RESULT 1
US-10-092-750-65
; Sequence 65, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/05002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR FILING DATE: 2001-03-08
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-65

Query Match 100.0%; Score 69; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SGLRKPTCGSSQR 13
Db 1 SGLRKPTCGSSQR 13
RESULT 2
US-10-437-963-138965
; Sequence 138965, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:

```

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 138965
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(323)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_40303C.1.pap
; US-10-437-963-138965

```

```

Query Match      71.0%; Score 49; DB 16; Length 95;
Best Local Similarity 66.7%; Pred. No. 2.9;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 GLRKPTCGSSOR 13
Db      2 GTRKPTCGEKQK 13

```

```

RESULT 3
US-10-425-115-300236
; Sequence 300236, Application US/10425115
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 300236
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_36886C.1.pap
; US-10-425-115-300236

```

```

Query Match      62.3%; Score 43; DB 17; Length 53;
Best Local Similarity 61.5%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      1 SGLRKPTCGSSOR 13
Db      41 SALEKPTCGSDR 53

```

```

RESULT 4
US-10-437-963-109125
; Sequence 109125, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei

```

```

; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 109125
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(323)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13312C.1.pap
; US-10-437-963-109125

```

```

Query Match      59.4%; Score 41; DB 16; Length 323;
Best Local Similarity 61.5%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      1 SGLRKPTCGSSOR 13
Db      195 SATRRPTAGSSSR 207

```

```

RESULT 5
US-10-437-963-138962
; Sequence 138962, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 138962
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(639)
; OTHER INFORMATION: Clone ID: PAT_MRT4530_40300C.1.pap
; US-10-437-963-138962

```

```

Query Match      59.4%; Score 41; DB 16; Length 639;
Best Local Similarity 58.3%; Pred. No. 3.2e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      2 GLRKPTCGSSOR 13
Db      470 GRDKPTCGEKQK 481

```

```

RESULT 6
US-10-425-115-312228
; Sequence 312228, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

```

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 312228
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_47809C.1.pbp
US-10-425-115-312228

Query Match 58.0%; Score 40; DB 17; Length 90;
Best Local Similarity 56.7%; Pred. No. 77;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GLRKPTCGSSQR 13
Db 46 GLRVPLCWGSQR 57

RESULT 7
US-10-425-115-338616
; Sequence 338616, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 338616
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(103)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_71985C.1.pbp
US-10-425-115-338616

Query Match 58.0%; Score 40; DB 17; Length 103;
Best Local Similarity 75.0%; Pred. No. 87;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GLRKPTCG 9
Db 55 GLRVPLCWGSQR 62

RESULT 8
US-10-437-963-161045
; Sequence 161045, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 161045
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(137)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_60266C.1.pbp
US-10-437-963-161045

Query Match 58.0%; Score 40; DB 16; Length 137;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSQ 12
Db 17 AGLKPLCASFQ 28

RESULT 9
US-10-425-114-45264
; Sequence 45264, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 45264
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 701181349_FLI.pbp
US-10-425-114-45264

Query Match 58.0%; Score 40; DB 15; Length 245;
Best Local Similarity 63.6%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSS 11
Db 87 SGRTEPCGAS 97

RESULT 10
US-10-425-115-198860
; Sequence 198860, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 198860
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_112937C.1.pap
US-10-425-115-198860

Query Match      58.0%; Score 40; DB 17; Length 291;
Best Local Similarity 63.6%; Pred. No. 2.2e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GLRKPTCGSSQ 12
DB      228 GWRCTCGAAQ 238

RESULT 11
US-10-425-115-256799
; Sequence 256799, Application US/10425115
; Publication No. US2004014272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 256799
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(383)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_16578C.1.pap
US-10-425-115-256799

Query Match      58.0%; Score 40; DB 17; Length 383;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 SGLRKPTCGS 10
DB      344 SGRRLPTCGS 353

RESULT 12
US-10-437-963-122034
; Sequence 122034, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Bockharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
```

```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 122034
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_24C.1.pap
US-10-437-963-122034

Query Match      58.0%; Score 40; DB 16; Length 734;
Best Local Similarity 53.8%; Pred. No. 5.2e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 SGLRKPTCGSSOR 13
DB      303 TGVRRPSSASSOR 315

RESULT 13
US-10-094-749-2840
; Sequence 2840, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMANOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOKIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2840
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2840

Query Match      58.0%; Score 40; DB 14; Length 740;
Best Local Similarity 61.5%; Pred. No. 5.2e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 SGLRKPTCGSSOR 13
DB      249 SGRFKPTKNSER 261

RESULT 14
US-10-408-765A-1557
; Sequence 1557, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylox, Steven W.
; APPLICANT: Glenn, Gary W.
; APPLICANT: Watnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1557
; LENGTH: 1011
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1557

Query Match      58.0%; Score 40; DB 16; Length 1011;
Best Local Similarity 87.5%; Pred. No. 7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GLRKPTCG 9
      |||||
DB      462 GLLKPTCG 469

RESULT 15
US-10-177-293-332
; Sequence 332, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jk., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Fuszta, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 1225
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-332

Query Match      58.0%; Score 40; DB 14; Length 1225;
Best Local Similarity 87.5%; Pred. No. 8.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GLRKPTCG 9
      |||||
DB      615 GLLKPTCG 622

Search completed: November 11, 2004, 07:41:45
Job time : 32.4128 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:52:32 ; Search time 4.44128 Seconds
(without alignments)
281.634 Million cell updates/sec

Title: US-10-092-750-65

Perfect score: 69

Sequence: 1 SGLRKPTCGSSQR 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR.79.*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45.5	65.9	438	T45785	hypothetical prote
2	43	62.3	212	G69369	branched-chain ami
3	42	60.9	476	C85041	probable DNA-bind
4	40	58.0	497	B42827	interleukin enhanc
5	40	58.0	816	A86266	hypothetical prote
6	40	58.0	2212	A41098	calcium channel pr
7	39	56.5	460	T45968	hypothetical prote
8	39	56.5	599	T18316	hypothetical prote
9	38	55.1	109	C72588	hypothetical prote
10	38	55.1	264	G83245	probable ATP-bind
11	38	55.1	386	AC0336	probable aminotran
12	38	55.1	386	T47022	hypothetical prote
13	38	55.1	619	T26453	hypothetical prote
14	37	53.6	195	S77401	hypothetical prote
15	37	53.6	376	JC4892	L-selectin precurs
16	37	53.6	385	T29315	hypothetical prote
17	37	53.6	440	T24323	hypothetical prote
18	37	53.6	1299	T158401	protein-tyrosine k
19	36.5	52.9	61	T1E8H	proteinase inhibit
20	36	52.2	51	G61588	ovomucoid (PSTI-ty
21	36	52.2	109	T49831	hypothetical prote
22	36	52.2	122	S56310	probable membrane
23	36	52.2	194	T16395	hypothetical prote
24	36	52.2	195	H96586	hypothetical prote
25	36	52.2	211	F71111	hypothetical prote
26	36	52.2	332	A87644	hypothetical prote
27	36	52.2	354	C71368	probable UDP-N-ace
28	36	52.2	373	A55718	interleukin-2 rece
29	36	52.2	404	B70522	hypothetical prote

30 36 52.2 447 2 T49439 hypothetical prote
31 36 52.2 478 2 AC3412 dihydroorotase (EC
32 36 52.2 559 2 E84213 hypothetical prote
33 36 52.2 585 2 S55205 dihydroxy-acid deh
34 36 52.2 780 2 G84708 probable VP1/AB13
35 36 52.2 899 2 B87553 DNA topoisomerase
36 36 52.2 1128 1 Q9BE47 DNA-binding protei
37 36 52.2 1590 2 B87754 protein C33E11.3 l
38 36 52.2 1705 2 S51672 adenylate cyclase
39 36 52.2 1706 1 OVBRC cyclolysin - Borda
40 35.5 51.4 162 2 D96581 hypothetical prote
41 35.5 51.4 213 2 F70080 hypothetical prote
42 35 50.7 80 2 AF3364 hypothetical prote
43 35 50.7 111 2 S25343 QR15 protein - yea
44 35 50.7 135 2 AB2373 mannose-6-phosphat
45 35 50.7 161 2 JC4728 mob protein B - Er

ALIGNMENTS

RESULT 1

T45785
hypothetical protein F26013.180 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T45785
R:Idelsen, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23013
A:Accession: T45785
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-438
A:Cross-references: UNIPROT:Q9SCZ5; EMBL:AL133452
A:Experimental source: cultivar Columbia; BAC clone F26013
C:Genetics:
A:Map position: 3
A:Introns: 35/1; 67/2; 104/3; 326/3
A>Note: F26013.180

Query Match 65.9%; Score 45.5; DB 2; Length 438;
Best Local Similarity 83.3%; Pred. No. 3.3;
Matches 10; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 SGLRKPTCGSSQ 12
Db 354 SGLRRP-COSSE 364
|||||
|||||

RESULT 2

G69369
branched-chain amino acid ABC transporter, ATP-binding protein (braf-3) homolog - Archa
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: G69369
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodsc
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiaich, P.; Kaine, B.P.; Sykes, S
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: G69369
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-212 <KLS>
A:Cross-references: UNIPROT:Q29303; GB:AE001038; GB:AE000782; NID:G2689361; PIDN:AB9302
C:Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
F:17-207/Domain: ATP-binding cassette homology <ABC>
F:32-39/Region: nucleotide-binding motif A (P-loop)

```

Query Match      62.3%; Score 43; DB 2; Length 212;
Best Local Similarity 77.8%; Pred. No. 5;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 SGLRPTCG 9
   ||| |||||
Db 46 AGLKPTCG 54

RESULT 3
CS5041
probable DNA-binding protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: C85041
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: C85041
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-476 <STO>
A:Cross-references: UNIPROT:Q9ZR06; GB:NC_001268; NID:g7270195; PIDN:CAB77810.1; GSPDB:G
A:Gene: AT4G03250
A:Map position: 4

Query Match      60.9%; Score 42; DB 2; Length 476;
Best Local Similarity 53.8%; Pred. No. 14;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGLRPTCGSSQR 13
   ||| |||||
Db 69 SGLRQDSGSTRK 81

RESULT 4
B4287
interleukin enhancer-binding factor ILF-2 - human
C:Species: Homo sapiens (man)
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 12-Sep-1997
C:Accession: B4287
R:Li, C.; Lusis, A.J.; Sparkes, R.; Nirula, A.; Gaynor, R.
Genomics 13, 665-671, 1992
A:Title: Characterization and chromosomal mapping of the gene encoding the cellular DNA
A:Reference number: A42827; MUID:92347863; PMID:1339390
A:Accession: B4287
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-497 <LII>
A>Note: sequence inconsistent with the nucleotide translation
A>Note: sequence extracted from NCBI backbone (NCBIN:109781, NCBIPI:109782)
C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
F:141-233/Domain: fork head DNA-binding domain homology <FHD>

Query Match      58.0%; Score 40; DB 2; Length 497;
Best Local Similarity 70.0%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GLRKPTCGSS 11
   ||| |||||
Db 482 GLRRPFCASS 491

RESULT 5
A86266
hypothetical protein P3F19.19 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86266
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86266
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-816 <STO>
A:Cross-references: UNIPROT:Q9SAF0; GB:AE005172; NID:G4850400; PIDN:AAD31070.1; GSPDB:G
C:Genetics:
C:Superfamily: oxysterol-binding protein; pleckstrin repeat homology

Query Match      58.0%; Score 40; DB 2; Length 816;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LRKPTCGSSQR 13
   ||| |||||
Db 153 LRKASCGSSNR 163

RESULT 6
A41098
calcium channel protein alpha-1 chain isoform A - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 09-Jul-2004
C:Accession: A41098; A35901
R:Starr, T.V.B.; Prystay, W.; Snutch, T.P.
Proc. Natl. Acad. Sci. U.S.A. 88, 5621-5625, 1991
A:Title: Primary structure of a calcium channel that is highly expressed in the rat cer
A:Reference number: A41098; MUID:91288516; PMID:1648226
A:Accession: A41098
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2212 <STA>
A:Cross-references: UNIPROT:P54282; GB:M64373; NID:g203110; PIDN:AAA40806.1; PID:g20311
R:Snutch, T.P.; Leonard, J.P.; Gilbert, M.M.; Lester, H.A.; Davidson, N.
Proc. Natl. Acad. Sci. U.S.A. 87, 3391-3395, 1990
A:Title: Rat brain expresses a heterogeneous family of calcium channels.
A:Reference number: A35901; MUID:90239020; PMID:1692134
A:Accession: A35901
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr
A:Molecule type: mRNA
A:Residues: 1435-1667 <SNU>
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match      58.0%; Score 40; DB 2; Length 2212;
Best Local Similarity 60.0%; Pred. No. 11e-02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGLRPTCGS 10
   ||| |||||
Db 1726 SGIQKPECN 1735

RESULT 7
T45968
hypothetical protein F7J8.260 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T45968
R;Sevan, M.; Zimmermann, W.; Gruenewald, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Le
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23018
A:Accession: T45968
A>Status: preliminary

```

```
A;Molecule type: DNA
A;Residues: 1-460 <BEV>
A;Cross-references: UNIPROT:Q9LFA8; EMBL:AL137189
A;Experimental source: cultivar Columbia; BAC clone F7J8
C;Genetics:
A;Map position: 5
A;Introns: 28/2; 57/3; 398/3
A;Note: F7J8.260

Query Match      56.5%; Score 39; DB 2; Length 460;
Best Local Similarity 53.8%; Pred. No. 46;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SGLRKPTCGSSQR 13
   |||:|:|
Db 80 SGIRRPSSSSSR 92

RESULT 8
T18316
hypothetical protein L7610.6 - Leishmania major
C;Species: Leishmania major
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18316
R;Oliver, K.; Murphy, L.; Quail, M.; Lawson, D.; Harris, D.; Rajadream, M.; Ivens, A.;
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18876
A;Accession: T18316
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-599 <OLI>
A;Cross-references: UNIPROT:Q97009; EMBL:AL034356; NID:el1371878; PID:el1371561; PIDN:CA42
C;Genetics:
A;Note: L7610.6

Query Match      56.5%; Score 39; DB 2; Length 599;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SGLRKPTCG 9
   |||:|:|
Db 63 TGLRRPQCG 71

RESULT 9
C72588
hypothetical protein APE1176 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: C72588
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: C72588
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-109 <KAW>
A;Cross-references: UNIPROT:Q9YCT6; DDBJ:AP000061; NID:95104821; PIDN:BA80161.1; PID:cl
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1176

Query Match      55.1%; Score 38; DB 2; Length 109;
Best Local Similarity 63.6%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LRKPTCGSSQR 13
   |||:|:|
Db 47 IRRETPGSSKR 57

RESULT 10
G83245
probable ATP-binding component of ABC transporter PA3212 [imported] - Pseudomonas aeru
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: G83245
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; I
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pat
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: G83245
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-264 <STO>
A;Cross-references: UNIPROT:Q9HZ28; GB:AE004744; GB:AE004091; NID:g9949317; PIDN:AA061
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA3212

Query Match      55.1%; Score 38; DB 2; Length 264;
Best Local Similarity 77.8%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GLRKPTCGS 10
   |||:|:|
Db 57 GLRPTSGS 65

RESULT 11
AC0236
probable aminotransferase YP01936 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0236
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.E
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 415, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0236
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-386 <KUR>
A;Cross-references: UNIPROT:Q9ZC46; GB:AL590842; PIDN:CAC90751.1; PID:g15979953; GSPDB:
C;Genetics:
A;Gene: YP01936
C;Superfamily: Escherichia coli valine-pyruvate transaminase

Query Match      55.1%; Score 38; DB 2; Length 386;
Best Local Similarity 70.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RXPTCGSSQR 13
   |||:|:|
Db 208 RXPVCASLQR 217

RESULT 12
T47022
hypothetical protein [imported] - Yersinia pestis
C;Species: Yersinia pestis
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C;Accession: T47022
R;Buchrieser, C.; Rusniok, C.; Couve, E.; Frangeul, L.; Billault, A.; Kunst, F.; Carnie
submitted to the EMBL Data Library, October 1998
A;Description: DNA sequence of the 102 kbases unstable region of Yersinia pestis.
A;Reference number: Z24348
A;Accession: T47022
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
```

A:Residues: 1-386 <BUC>
A:Cross-References: UNIPROT:Q9ZC46; EMBL:AL031866; PIDN:CAA21365.1
A:Experimental source: strain 6/69
C:Superfamily: Escherichia coli valine-pyruvate transaminase

Query Match 55.1%; Score 38; DB 2; Length 386;
Best Local Similarity 70.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 RKPTCGSSQR 13
||| ||| |||
DB 208 RKPVCASLQR 217

RESULT 13

T26453
Hypothetical protein ZC15.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26453; T27501
R:Lennard, N.
submitted to the EMBL Data Library, November 1998

A:Reference number: Z20216

A:Accession: T26453

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-619 <WIL>

A:Cross-References: UNIPROT:O18266; EMBL:AL033509; PIDN:CAA22060.1; CESP:ZC15.2

A:Experimental source: clone Y113G7C

R:Basham, V.
submitted to the EMBL Data Library, March 1997

A:Reference number: Z20377

A:Accession: T27501

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-619 <W12>

A:Cross-References: EMBL:Z93396; NID:e1188462; PIDN:CAB07710.1; GSPDB:GN00023; CESP:ZC15

A:Experimental source: clone ZC15

C:Genetics:

A:Gene: CESP:ZC15.2

A:Map position: 5

A:Introns: 31/2; 216/3; 279/3; 343/3; 400/3; 428/1; 450/3; 489/1; 532/3; 593/2

Query Match 55.1%; Score 38; DB 2; Length 619;
Best Local Similarity 70.0%; Pred. No. 86;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGLRKPTCGS 10
||| ||| |||
DB 508 SGMKXPECYS 517

RESULT 14

S77401

Hypothetical protein sir1222 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C:Accession: S77401

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

s.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S77401

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-195 <KAN>

A:Cross-References: UNIPROT:P73464; EMBL:D90906; GB:AB001339; NID:G1652492; PIDN:BAAL1750

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: Synechocystis hypothetical protein sir1222

Query Match 53.6%; Score 37; DB 2; Length 195;
Best Local Similarity 54.5%; Pred. No. 50;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LRKPTCGSSQR 13
||| ||| |||
DB 169 LRRPTCAFEQK 179

RESULT 15

JC4892

L-selectin precursor - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004

C:Accession: JC4892

R:Qian, J.; Huang, X.; Marks, R.M.

Biochem. Biophys. Res. Commun. 225, 406-412, 1996

A:Title: Cloning of the cDNA for rabbit L-selectin and expression of recombinant protein

A:Reference number: JC4892; MUID:96354800; PMID:8753776

A:Accession: JC4892

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-376 <QIA>

A:Cross-References: UNIPROT:Q28629; GB:U26535; NID:G847787; PIDN:AAA67896.1; PID:G84778

C:Comment: This protein involved in leukocyte-endothelial adhesion; it mediates adhesion

C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;

F;1-37/Domain: signal sequence #status predicted <SIG>

F;29-155/Domain: C-type lectin homology <LCH>

F;38-376/Product: L-selectin #status predicted <MAT>

F;160-191/Domain: EGF homology <EGF>

F;197-254/Domain: complement factor H repeat homology <FHL>

F;259-316/Domain: complement factor H repeat homology <FHR>

Query Match 53.6%; Score 37; DB 2; Length 376;
Best Local Similarity 60.0%; Pred. No. 85;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GLRKPTCGSS 11
||| ||| |||
DB 297 GMKKTVCGSS 306

Search completed: November 10, 2004, 15:55:03
Job time: 5.44128 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 21.79 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-65
Perfect score: 69
Sequence: 1 SGLRKPTCGSSQR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 02.*
1: uniprot_sprot.*
2: uniprot_tramb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	72.5	262	Q88K75	Q88K75 pseudomonas
2	45.5	65.9	438	Q9SC25	Q9SC25 arabidopsis
3	43	62.3	212	Q29303	Q29303 archaeoglob
4	43	62.3	354	Q6YI28	Q6YI28 oryza sativ
5	43	62.3	354	QAD17472	QAD17472 oryza sat
6	42	60.9	476	Q9ZRO6	Q9ZRO6 arabidopsis
7	42	60.9	548	Q880R1	Q880R1 pseudomonas
8	42	60.9	733	Q95KL5	Q95KL5 macaca fasc
9	41	59.4	274	Q9XH68	Q9XH68 brassica na
10	41	59.4	1044	Q6WQJ1	Q6WQJ1 mus muscu
11	41	59.4	1044	Q6WQJ1	Q6WQJ1 mus muscu
12	40	58.0	65	1 ICHY_CAIMO	1 ICHY_CAIMO
13	40	58.0	117	Q8VMG9	Q8VMG9 pseudomonas
14	40	58.0	321	Q872H3	Q872H3 pseudomonas
15	40	58.0	325	Q9DG19	Q9DG19 oncorhynch
16	40	58.0	329	Q8PYR6	Q8PYR6 methanocarc
17	40	58.0	334	Q7S248	Q7S248 neurospora
18	40	58.0	462	Q6NR48	Q6NR48 neurospora
19	40	58.0	462	Q6NR48	Q6NR48 neurospora
20	40	58.0	550	Q70369	Q70369 xenopus lae
21	40	58.0	593	Q6DU55	Q6DU55 xenopus lae
22	40	58.0	740	Q96M11	Q96M11 homo sapien
23	40	58.0	791	Q70368	Q70368 rattus norv
24	40	58.0	816	Q70368	Q70368 rattus norv
25	40	58.0	1011	Q3BXX2	Q3BXX2 arabidopsis
26	40	58.0	2212	1 CCAA_RAT	1 CCAA_RAT
27	39	56.5	99	Q6X631	Q6X631 r voltage-d
28	39	56.5	99	Q6X631	Q6X631 ancylostoma
29	39	56.5	99	CAE14754	CAE14754 leptospir
30	39	56.5	148	Q9P084	Q9P084 ancylostom
31	39	56.5	202	Q8P6H0	Q8P6H0 xanthomonas

Q8dyg7 streptococc
Q8e430 streptococc
Q7urhl rhodopirell
Q7rsi8 neurospora
Q7nez5 gloeobacter
Q6dcs2 xenopus lae
Q9lfa8 arabidopsis
Q97009 leishmania
Q6nx15 mus musculu
Aah67016 mus musculu
Q88143 chimpanzee
Q9w343 drosophila
Q6ddj0 xenopus lae
Q8r0x3 mus musculu

ALIGNMENTS

RESULT 1

Q88K75 PRELIMINARY; PRT; 262 AA.
ID Q88K75
AC Q88K75
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Iron ABC transporter, ATP-binding protein, putative.
GN OrderedLocusNames=PP2416;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzez A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Dueterhoeft A., Tuemmler B.,
RA Frazer C.M.;
RT "Complete genome sequence and comparative analysis of the
metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AR016783; AAN68028.1; -.
DR HSSP; P06611; 1L7V.
DR TIGR; PP2416; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00005; ABC_tran; 1.
DR Prodom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
SQ ATP-binding; Complete proteome.
KW SEQUENCE 262 AA; 28148 MW; AE1799A531CD7073 CRC64;

Query Match 72.5%; Score 50; DB 2; Length 262;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 SGLRKPTCGSSQ 12
DB 58 AGLRXPACGSVQ 69

```

RESULT 2
ID Q9SCZ5 PRELIMINARY; PRT; 438 AA.
AC Q9SCZ5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F26013.180.
GN Name=F26013.180;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RA Delsenly M., Berger C., Cooke R., Grellet F., Laudie M., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2];
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133452; CAB63018.1; -.
DR PIR; T45785; T45785.
KW Hypothetical protein.
SQ SEQUENCE 438 AA; 47254 MW; C0402B1D2F43CF0B CRC64;

Query Match 65.9%; Score 45.5; DB 2; Length 438;
Best Local Similarity 83.3%; Pred. No. 14;
Matches 10; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 SGLRPTCGSSQ 12
Db 354 SGLRPT-CGSSE 364
|||||

RESULT 3
ID O29303 PRELIMINARY; PRT; 212 AA.
AC O29303;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Branched-chain amino acid ABC transporter, ATP-binding protein (Braf-3) (Methanococcus jannaschii).
DE OrderedLocusNames=AF0959;
GN Archaeoglobus fulgidus.
OS Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475; DOI=10.1038/37052;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M.L., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,
RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodek A.,
RA Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A.,
RA Utterback T.R., Cotton M.D., Spriggs T., Artiach P., Kaine B.P.,
RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,
RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,
RA Woese C.R., Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF001038; AAB90284.1; -.
DR PIR; G69369; G69369.

```

```

DR HSSP; Q58663; 1G9X.
DR TIGR; AF0959; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00005; ABC tran; 1.
DR ProDom; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00111; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 212 AA; 23606 MW; 625F87A47803390C CRC64;

Query Match 62.3%; Score 43; DB 2; Length 212;
Best Local Similarity 77.8%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGLRPTCG 9
Db 46 AGLEKPTCG 54
|||||

RESULT 4
ID Q6YY28 PRELIMINARY; PRT; 354 AA.
AC Q6YY28;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNB0056122.17;
GN Names=OSJNB0056122.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1];
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005644; BAD17472.1; -.
KW Hypothetical protein.
SQ SEQUENCE 354 AA; 38565 MW; 745AB1B003575BEEA CRC64;

Query Match 62.3%; Score 43; DB 2; Length 354;
Best Local Similarity 70.0%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GLRKPTCGSS 11
Db 33 GMRAPSCGSS 42
|||||

RESULT 5
ID BAD17472 PRELIMINARY; PRT; 354 AA.
AC BAD17472;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNB0056122.17.
GN OSJNB0056122.17.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza; Oryza sativa.
OX NCBI_TaxID=39947;
RN [1];
RP SEQUENCE FROM N.A.

```

RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
 clone:OSUNB0056122.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005644; BADI17472.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 354 AA; 38565 MW; 745AB1B003575BEA CRC64;

Query Match 62.3%; Score 43; DB 2; Length 354;
 Best Local Similarity 70.0%; Pred. No. 32;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GLRKPTCGSS 11
 | | | | |
 Db 33 GWRAPSCSS 42

RESULT 6

Q9ZR06 PRELIMINARY; PRT; 476 AA.
 ID Q9ZR06
 AC Q9ZR06
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative DNA-binding protein (Homeodomain protein 14).
 GN Name=F4C21.18; Synonyms=At4g03250, hd-14;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;
 OC eurosidii; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 [1]
 RN SEQUENCE FROM N.A.
 RA Habermann K., de la Bastide M., Huang E.N., Gnoj L., Schutz K.,
 RA Preston R., Calma C., Martienssen R., Parnell L.D., Dedhia N.,
 RA McCombie W.R.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Parnell L.D., McCombie W.R.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Parnell L.D.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Ciabelli A.R., Carabelli M., Ruzza V., Sessa G., Steindler C.,
 RA Ruberti I.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
 RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
 RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
 RA Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 DR EMBL; AC005275; AAD14453.1; -.
 DR EMBL; AJ441296; CAD29664.1; -.
 DR EMBL; AL161496; CAB77810.1; -.
 DR FIR; C85041; C85041.
 DR HSSP; P10037; 1AU7.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001356; Homeobox.
 DR InterPro; IPR009057; Homeodomain_like.

DR Pfam; PF00046; Homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS0071; HOMEBOX_2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 SQ SEQUENCE 476 AA; 53493 MW; D9322A290AFC4F91 CRC64;

Query Match 60.9%; Score 42; DB 2; Length 476;
 Best Local Similarity 53.8%; Pred. No. 65;
 Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSOR 13
 | | | | |
 Db 69 SGLRQDSGSGTKQ 81

RESULT 7

Q880R1 PRELIMINARY; PRT; 548 AA.
 ID Q880R1
 AC Q880R1
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Nickel ABC transporter, ATP-binding protein, putative.
 GN OrderedLocusNames=PSPO3091;
 OS Pseudomonas syringae (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=323;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=DC3000;
 RC MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
 RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
 RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
 RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
 RA Nelson W.C., Daviden T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
 RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
 RA Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
 RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
 RA Bander C.L., White O., Fraser C.M., Collier A.;
 RT "The complete genome sequence of the Arabidopsis and tomato pathogen
 Pseudomonas syringae pv. tomato DC3000.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; AB016867; AAO56580.1; -.
 DR HSSP; Q58206; 1F30.
 DR TIGR; PSPT03091; -.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005224; F:ATP binding; IEA.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
 DR GO; GO:0000166; P:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR PROSITE; PS0211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
 KW ATP-binding; Complete proteome.
 SQ SEQUENCE 548 AA; 59616 MW; 38362F8DAB018794 CRC64;

Query Match 60.9%; Score 42; DB 2; Length 548;
 Best Local Similarity 87.5%; Pred. No. 75;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GLRKPTCG 9
 | | | | |
 Db 340 GLEKPTCG 347

```

RESULT 8
Q95K15          PRELIMINARY;      PRT; 733 AA.
AC Q95K15;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
DE Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Osada N., Hida M., Kusuda J., Tanuma R., Hirata M., Suto Y., Hirai M.,
RA Terao K., Sugano S., Hashimoto K.;
RT "Cynomolgus monkey testicular cDNAs for discovery of novel human genes
RT in the human genome sequence.";
RL BMC Genomics 3:36-36(2002).
DR EMBL; AB070012; BAB62957.1; -.
DR HSSP; Q00420; LAMC.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank. 3.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS00888; ANK_REPEAT; 3.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Hypothetical protein.
SQ SEQUENCE 733 AA; 81674 MW; 0D4E69C1B8E025AE CRC64;

Query Match 60.9%; Score 42; DB 2; Length 733;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSSQ 12
Db 330 SGLSPVCGSMQ 341

RESULT 9
Q9XH68          PRELIMINARY;      PRT; 274 AA.
AC Q9XH68;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 33 kDa oxygen evolving protein of photosystem II.
GN Name=psbO;
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=canola;
RA Gao Y.P., Gusta L.V.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF139818; AAD38521.1; -.
DR InterPro; IPR002628; PSII_MSP.
DR Pfam; PF01716; MSP; 1.
SQ SEQUENCE 274 AA; 29642 MW; 357A3B621FF41A0D CRC64;

Query Match 59.4%; Score 41; DB 2; Length 274;
Best Local Similarity 70.0%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 RKPTCGSSQR 13
Db 205 RKPTCASSSR 214

```

```

RESULT 10
Q6WQJ1          PRELIMINARY;      PRT; 1044 AA.
AC Q6WQJ1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Neural stem cell-derived dendrite regulator.
GN Name=Nsddr;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Horiguchi S., Tashiro K., Takahashi J., Hashimoto N., Nakano I.,
RA Tsuchida Y., Hirai H., Honjo T.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY275376; AAQ17118.1; -.
DR InterPro; IPR002921; Lipase_3.
DR InterPro; IPR008262; Lipase_AS.
DR Pfam; PF01764; Lipase_3; 1.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN 1.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN 1.
SQ SEQUENCE 1044 AA; 115375 MW; E2AC2349D4E3503E CRC64;

Query Match 59.4%; Score 41; DB 2; Length 1044;
Best Local Similarity 70.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LRKPTCGSSQ 12
Db 310 MRKPTCGLCQ 319

RESULT 11
AAQ17118        PRELIMINARY;      PRT; 1044 AA.
AC AAQ17118;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Neural stem cell-derived dendrite regulator.
GN NSDDR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Horiguchi S., Tashiro K., Takahashi J., Hashimoto N., Nakano I.,
RA Tsuchida Y., Hirai H., Honjo T.;
RL "NSDDR a novel tetra-spanning transmembrane protein with a unique
RL integration pattern to the plasma membrane regulates the extension of
RL the dendritic trees of Purkinje cells.";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY275376; AAQ17118.1; -.
SQ SEQUENCE 1044 AA; 115375 MW; E2AC2349D4E3503E CRC64;

Query Match 59.4%; Score 41; DB 2; Length 1044;
Best Local Similarity 70.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LRKPTCGSSQ 12
Db 310 MRKPTCGLCQ 319

RESULT 12
ICHY_CAIMO

```



```

ID ICHY CAIMO STANDARD; PRT; 65 AA.
AC P83039;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chymotrypsin inhibitor (DPCI).
OS Cairina moschata (Muscovy duck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Cairina.
OX NCBI_TaxID=8855;
RN [1]
RP SEQUENCE, FUNCTION, SUBUNIT, AND MASS SPECTROMETRY.
RC TISSUE=Pancreas;
RX MEDLINE=21956294; PubMed=11959032;
RA Wilimowska-Pelc A., Olchiewicz Z., Mazurkiewicz A., Kowalska J.,
RA Wilusz T.;
RT "Kazal-type chymotrypsin inhibitor from duck pancreas.";
RL Comp. Biochem. Physiol. 131B:499-507(2002).
CC -!- FUNCTION: Inhibits chymotrypsin.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MASS SPECTROMETRY: MW=7190.9; MW_ERR=2; METHOD=MALDI; RANGE=1-65;
CC [1]
CC NOTE=Ref.1.
CC -!- SIMILARITY: Contains 1 Kazal-like domain.
DR HSSP; P00995; 1HPT.
DR InterPro; IPR002350; Prot_inh_Kazal.
DR InterPro; IPR001239; Prot_inh_Kazal-m.
DR Pfam; PF00050; Kazal; 1.
DR SMART; SM00290; KAZALINHSTR.
DR SMART; SM00280; KAZAL; 1.
DR PROSITE; PS00282; KAZAL; 1.
KW Direct protein sequencing; Serine protease inhibitor.
FT DOMAIN 11 62
FT 62 Kazal-like.
FT DISULFID 13 44
FT 44 By similarity.
FT DISULFID 22 41
FT 41 By similarity.
FT DISULFID 30 62
FT 62 By similarity.
FT SITE 24 25
FT 25 Reactive bond (By similarity).
SQ SEQUENCE 65 AA; 7195 MW; 360AA7E1AA515D47 CRC64;
Query Match 58.0%; Score 40; DB 1; Length 65;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGLRKPTCG 9
DB :|||:|
6 NGLRPVCG 14

RESULT 13
Q8VMG9 PRELIMINARY; PRT; 117 AA.
ID Q8VMG9;
AC Q8VMG9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Pseudomonas putida.
OC Plasmid pW0.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2243065; PubMed=12534468;
RA Great A., Lamberton L., Williams P.A., Thomas C.M.;
RT "Complete sequence of the IncP-9 TOL plasmid pW0 from Pseudomonas
RT putida.";
RL Environ. Microbiol. 4:856-871(2002).
DR EMBL; AJ344068; CAC86847.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 117 AA; 12700 MW; 22CBA6EEC067638D CRC64;
Query Match 58.0%; Score 40; DB 2; Length 117;

```

```

Best Local Similarity 77.8%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GLRKPTCGS 10
DB :|||:|
18 GLKPAACGS 26

RESULT 14
Q872H3 PRELIMINARY; PRT; 321 AA.
ID Q872H3;
AC Q872H3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Related to transcriptional activator CMR1.
GN Name=80A10.010;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Kannhaup G.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Contains 1 Zn(2)-Cys(6) fungal-type binuclear cluster
CC domain.
CC EMBL; BX294012; CAD70758.1; -.
DR HSSP; P07248; 2ADR.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0008270; P:zinc ion binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001138; Fungi_Trcrp_N.
DR InterPro; IPR007087; ZnF_C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR Pfam; PF00172; Zn clus; 1.
DR PRINTS; PR00054; FUNGALZNCYS.
DR SMART; SM00066; GAL4; 1.
DR SMART; SM00355; ZNF_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
DR PROSITE; PS00463; ZN2_CY6_FUNGAL_1; UNKNOWN_1.
DR PROSITE; PS00048; ZN2_CY6_FUNGAL_2; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Transcription regulation;
KW Zinc.
SQ SEQUENCE 321 AA; 35330 MW; 91E67B69213B9E9C CRC64;
Query Match 58.0%; Score 40; DB 2; Length 321;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 RKPTCGSSQR 13
DB :|||:|:|
91 QRPTCGACQK 100

RESULT 15
Q9DG19 PRELIMINARY; PRT; 325 AA.
ID Q9DG19;
AC Q9DG19;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DMRT1 protein.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_taxid=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20435294; PubMed=10978520;
 RA Marchand O., Govoroun M., D'Cotta H., McMeel O., Lareyre J.J.,
 RA Bernot A., Laudet V., Guiguen Y.;
 RT "DMRT1 expression during gonadal differentiation and spermatogenesis
 RT in the rainbow trout, *Oncorhynchus mykiss*.";
 RL Biochim. Biophys. Acta 1493:180-187(2000).
 DR EMBL; AF209095; AAG17544.1; -.
 DR HSSP; P23023; LEV.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007548; P:sex differentiation; IEA.
 DR InterPro; IPR001275; DM_DNA-binding.
 DR Pfam; PF00751; DM; 1.
 DR SMART; SM00301; DM; 1.
 DR PROSITE; PS40000; DM_DOMAIN_1; 1.
 DR PROSITE; PS50809; DM_DOMAIN_2; 1.
 SQ SEQUENCE 325 AA; 34613 MW; 4FPD2A710F9DF431 CRC64;

Query Match 58.0%; Score 40; DB 2; Length 325;
 Best Local Similarity 63.6%; Pred. NO. 1e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGLRKPTCGSS 11
 Db 113 SGRSPTCGNT 123

Search completed: November 10, 2004, 15:53:24
 Job time : 23.867 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 11.5658 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-66
Perfect score: 132
Sequence: 1 AGTQPLILAQFMVGGDELLHFLW 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50.5	38.3	642	4	US-09-252-991A-27772
2	50	37.9	62	3	US-09-537-357-11
3	50	37.9	480	3	US-09-078-173A-25
4	50	37.9	480	3	US-09-537-357-32
5	50	37.9	480	4	US-10-042-991-25
6	48	36.4	149	4	US-09-583-110-3795
7	48	36.4	219	4	US-09-270-767-46059
8	48	36.4	1036	2	US-08-720-48A-5
9	48	36.4	1036	3	US-08-953-823A-5
10	48	36.4	1036	4	US-09-398-239-5
11	48	36.4	1036	4	US-09-560-876A-5
12	48	36.4	1065	4	US-09-560-876A-6
13	47.5	36.0	341	4	US-09-252-991A-18920
14	46	34.8	399	4	US-09-252-991A-19103
15	46	34.8	566	4	US-09-252-991A-26648
16	46	34.8	957	4	US-09-252-991A-24683
17	45	34.1	316	4	US-09-252-991A-32914
18	45	34.1	597	4	US-09-328-352-4703
19	45	34.1	916	4	US-09-248-796A-25756
20	44.5	33.7	812	4	US-09-489-039A-12075
21	44	33.3	316	4	US-09-252-991A-20585
22	43.5	33.0	251	4	US-09-270-767-41768
23	43.5	33.0	1304	4	US-09-489-039A-13449
24	43	32.6	72	4	US-09-206-551-36
25	43	32.6	387	4	US-09-252-991A-21543
26	43	32.6	401	4	US-09-134-000C-4410
27	43	32.6	420	4	US-09-328-352-7296

Sequence 2, Appl
Sequence 8223, Ap
Sequence 28693, A
Sequence 22239, A
Sequence 21996, A
Sequence 7532, Ap
Sequence 334, App
Sequence 4674, Ap
Sequence 11898, A
Sequence 17172, A
Sequence 19064, A
Sequence 31999, A
Sequence 11865, A
Sequence 32609, A
Sequence 41, Appl
Sequence 4955, Ap
Sequence 2052, Ap
Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-252-991A-27772
; Sequence 27772, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27772
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27772

Query Match 38.3%; Score 50.5; DB 4; Length 642;
Best Local Similarity 42.9%; Pred. No. 6.5;
Matches 12; Conservative 4; Mismatches 9; Indels 3; Gaps 1;
QY 1 AGTQPLILAQFMVGGDELLH---FLW 25
DB 311 AGNPAIGDRLEVGDDVAARRVFLW 338

RESULT 2
US-09-537-357-11
; Sequence 11, Application US/09537357
; Patent No. 6271018
; GENERAL INFORMATION:
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: MUSKELON (CUCUMIS MELO) HYDROPEROXIDE
; TITLE OF INVENTION: LYASE AND USES THEREOF
; FILE REFERENCE: 06027.0002
; CURRENT APPLICATION NUMBER: US/09/537,357
; CURRENT FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-09-537-357-11

```
Query Match      37.9%; Score 50; DB 3; Length 62;
Best Local Similarity 31.6%; Pred. No. 0.53;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY  2 GTQPLI-----LAQFMRVGGDELLHFLW 25
Db   6 GYQPLWKDPKVFDEPEKFMLEFRTKEGKELLNLFW 43

RESULT 3
US-09-078-173A-25
; Sequence 25, Application US/09078173A
; Patent No. 6200794
; GENERAL INFORMATION:
; APPLICANT: Ian M. Whitehead
; APPLICANT: Alan Slusarenko
; APPLICANT: Duncan Gaskin
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: GUAVA (PSIDIUM GUAJAVA) 13-HYDROPEROXIDE
; FILE REFERENCE: 06027.0001
; CURRENT APPLICATION NUMBER: US/09/078,173A
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-09-078-173A-25

Query Match      37.9%; Score 50; DB 3; Length 480;
Best Local Similarity 31.6%; Pred. No. 5.6;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY  2 GTQPLI-----LAQFMRVGGDELLHFLW 25
Db   389 GYQPLWKDPKVFDEPEKFMLEFRTKEGKELLNLFW 426

RESULT 4
US-09-537-357-32
; Sequence 32, Application US/09537357
; Patent No. 6271018
; GENERAL INFORMATION:
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: MUSKMELON (CUCUMIS MELO) HYDROPEROXIDE
; FILE REFERENCE: 06027.0002
; CURRENT APPLICATION NUMBER: US/09/537,357
; CURRENT FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-09-537-357-32

Query Match      37.9%; Score 50; DB 3; Length 480;
Best Local Similarity 31.6%; Pred. No. 5.6;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY  2 GTQPLI-----LAQFMRVGGDELLHFLW 25
Db   389 GYQPLWKDPKVFDEPEKFMLEFRTKEGKELLNLFW 426

RESULT 5
US-10-042-991-25
```

```
; Sequence 25, Application US/10042991
; Patent No. 6780621
; GENERAL INFORMATION:
; APPLICANT: Ian M. Whitehead
; APPLICANT: Alan Slusarenko
; APPLICANT: Duncan Gaskin
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: GUAVA (PSIDIUM GUAJAVA) 13-HYDROPEROXIDE
; FILE REFERENCE: 06027.0001U3
; CURRENT APPLICATION NUMBER: US/10/042,991
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-10-042-991-25

Query Match      37.9%; Score 50; DB 4; Length 480;
Best Local Similarity 31.6%; Pred. No. 5.6;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY  2 GTQPLI-----LAQFMRVGGDELLHFLW 25
Db   389 GYQPLWKDPKVFDEPEKFMLEFRTKEGKELLNLFW 426

RESULT 6
US-09-583-110-2795
; Sequence 2795, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2795
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2795

Query Match      36.4%; Score 48; DB 4; Length 149;
Best Local Similarity 53.3%; Pred. No. 3.1;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY  11 FMRVGGDELLHFLW 25
Db   131 FLAQAQSDQLLHFLW 145

RESULT 7
US-09-270-767-46059
; Sequence 46059, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

GENERAL INFORMATION:
APPLICANT: DeSavage, Frederic
APPLICANT: Rosenthal, Arnon

```

RESULT 11
US-09-560-876A-5
; Sequence 5, Application US/09560876A
; Patent No. 6492139
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Rosenthal, Arnon
; APPLICANT: Stone, Donna
; TITLE OF INVENTION: Vertebrate Smoothened Proteins
; FILE REFERENCE: P1050R1D1
; CURRENT APPLICATION NUMBER: US/09/560,876A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/027,070
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: US 08/953,823
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-560-876A-5

Query Match      36.4%; Score 48; DB 4; Length 1036;
Best Local Similarity 43.5%; Pred. No. 29;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY      1 AGTQPLILAQFMRVGGDELLHFL 23
Db      788 AGLDPSILNEFLQKNGDFIFPFL 810

RESULT 12
US-09-560-876A-6
; Sequence 6, Application US/09560876A
; Patent No. 6492139
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Rosenthal, Arnon
; APPLICANT: Stone, Donna
; TITLE OF INVENTION: Vertebrate Smoothened Proteins
; FILE REFERENCE: P1050R1D1
; CURRENT APPLICATION NUMBER: US/09/560,876A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/027,070
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: US 08/953,823
; PRIOR FILING DATE: 1997-09-30
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 6
; LENGTH: 1065
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 3
; OTHER INFORMATION: unknown amino acid
US-09-560-876A-6

Query Match      36.4%; Score 48; DB 4; Length 1065;
Best Local Similarity 43.5%; Pred. No. 30;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY      1 AGTQPLILAQFMRVGGDELLHFL 23
Db      817 AGLDPSILNEFLQKNGDFIFPFL 839

RESULT 13
US-09-252-991A-18920
; Sequence 18920, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18920
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18920

Query Match      36.0%; Score 47.5; DB 4; Length 341;
Best Local Similarity 43.3%; Pred. No. 9.7;
Matches 13; Conservative 1; Mismatches 7; Indels 9; Gaps 1;

QY      4 QPLIL-----AQFMRVGGDELLHFL 24
Db      196 QFLVAQPRDPLREARQQRVGGELQHLAL 225

RESULT 14
US-09-252-991A-19103
; Sequence 19103, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19103
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19103

Query Match      34.8%; Score 46; DB 4; Length 399;
Best Local Similarity 52.6%; Pred. No. 21;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY      4 QPLILAQFMRVGGDELLHFL 22
Db      175 QPLLGLFELAGCGQLLQF 193

RESULT 15
US-09-252-991A-26648
; Sequence 26648, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19103
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26648

Query Match      34.8%; Score 46; DB 4; Length 399;
Best Local Similarity 52.6%; Pred. No. 21;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY      4 QPLILAQFMRVGGDELLHFL 22
Db      175 QPLLGLFELAGCGQLLQF 193

```

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26648
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26648

Query Match 34.8%; Score 46; DB 4; Length 566;
Best Local Similarity 55.0%; Pred.No. 31;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 6 LILAQFMRVGGDELHFLW 25
| : | | | | | | | | | |
Db 339 LVAAQDERFGADRLIVDLW 358

Search completed: November 10, 2004, 15:57:20
Job time : 12.5658 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 : Search time 60.4093 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-66

Perfect score: 132

Sequence: 1 AGTQPLILAQFMRVGGDELLHFLW 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	132	100.0	25	14 US-10-092-750-66	Sequence 66, Appl
2	53.5	40.5	168	17 US-10-425-115-229512	Sequence 229512,
3	51	38.6	161	15 US-10-425-114-48127	Sequence 48127, A
4	51	38.6	163	15 US-10-425-114-40985	Sequence 40985, A
5	50	37.9	62	9 US-09-884-260A-11	Sequence 11, Appl
6	50	37.9	70	17 US-10-425-115-327061	Sequence 327061,
7	50	37.9	476	15 US-10-381-870-5	Sequence 5, Appl1
8	50	37.9	480	9 US-09-884-260A-32	Sequence 32, Appl
9	50	37.9	480	13 US-10-042-991-25	Sequence 25, Appl
10	48	36.4	103	16 US-10-767-701-55335	Sequence 55335, A
11	48	36.4	276	15 US-10-425-114-53219	Sequence 53219, A
12	48	36.4	287	15 US-10-425-114-67552	Sequence 67552, A
13	48	36.4	2447	14 US-10-190-115-28	Sequence 28, Appl

14	48	36.4	2447	15	US-10-369-072-28	Sequence 28, Appl
15	47	35.6	315	14	US-10-306-762-214	Sequence 214, App
16	47	35.6	625	15	US-10-282-122A-77067	Sequence 77067, A
17	46	34.8	119	15	US-10-424-599-236063	Sequence 236063,
18	46	34.8	664	15	US-10-424-599-164152	Sequence 164152,
19	45.5	34.5	177	15	US-10-425-114-46400	Sequence 46400, A
20	45.5	34.5	177	15	US-10-425-114-73004	Sequence 73004, A
21	45	34.1	112	16	US-10-437-963-177358	Sequence 177358,
22	45	34.1	298	15	US-10-424-599-261542	Sequence 261542,
23	45	34.1	447	16	US-10-437-963-144679	Sequence 144679,
24	45	34.1	451	14	US-10-369-493-14329	Sequence 14329, A
25	45	34.1	451	14	US-10-369-493-14988	Sequence 14988, A
26	45	34.1	502	14	US-10-156-761-11794	Sequence 11794, A
27	45	34.1	597	16	US-10-437-963-147272	Sequence 147272,
28	45	34.1	577	16	US-10-437-963-144680	Sequence 144680,
29	45	34.1	708	16	US-10-437-963-204013	Sequence 204013,
30	44.5	33.7	1112	14	US-10-128-714-8353	Sequence 8353, Ap
31	44	33.3	61	17	US-10-425-115-333760	Sequence 333760,
32	44	33.3	105	17	US-10-425-115-310963	Sequence 310963,
33	44	33.3	110	17	US-10-425-115-305271	Sequence 305271,
34	44	33.3	121	17	US-10-425-115-219262	Sequence 219262,
35	44	33.3	268	14	US-10-107-096-2	Sequence 2, Appli
36	44	33.3	310	15	US-10-424-599-195450	Sequence 195450,
37	44	33.3	980	15	US-10-211-462-145	Sequence 145, App
38	44	33.3	1104	14	US-10-128-714-3353	Sequence 3353, Ap
39	43.5	33.0	148	10	US-09-882-227-4	Sequence 4, Appli
40	43.5	33.0	200	16	US-10-437-963-174116	Sequence 174116,
41	43.5	33.0	260	15	US-10-282-122A-50190	Sequence 50190, A
42	43.5	33.0	261	14	US-10-369-493-10659	Sequence 10659, A
43	43.5	33.0	1171	15	US-10-282-122A-77970	Sequence 77970, A
44	43.5	33.0	1300	15	US-10-282-122A-60085	Sequence 60085, A
45	43	32.6	51	16	US-10-767-701-50265	Sequence 50265, A

ALIGNMENTS

RESULT 1
US-10-092-750-66
; Sequence 66, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-66

Query Match 100.0%; Score 132; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.2e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGTQPLILAQFMRVGGDELLHFLW 25
DB 1 AGTQPLILAQFMRVGGDELLHFLW 25
RESULT 2
US-10-425-115-229512
; Sequence 229512, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 229512
LENGTH: 168
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(168)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_140909C.1.pap
US-10-425-115-229512

Query Match 40.5%; Score 53.5; DB 17; Length 168;
Best Local Similarity 44.0%; Pred. No. 2.4;
Matches 11; Conservative 4; Mismatches 3; Indels 7; Gaps 1;

QY 1 AGTQPLILAQFMRVGGDELLHFLW 25
DB 81 AGIMPLVLS-----EAVLHFLW 98

RESULT 3

US-10-425-114-48127
Sequence 48127, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 48127
LENGTH: 161
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3279-043-H12_FLI.pap
US-10-425-114-48127

Query Match 38.6%; Score 51; DB 15; Length 161;
Best Local Similarity 50.0%; Pred. No. 5.8;
Matches 10; Conservative 5; Mismatches 5; Indels 5; Gaps 0;

QY 1 AGTQPLILAQFMRVGGDELL 20
DB 22 AGLEPLLKARFGLAGGDL 41

RESULT 4

US-10-425-114-40985
Sequence 40985, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 40985
LENGTH: 163
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3059-042-B9_FLI.pap
US-10-425-114-40985

Query Match 38.6%; Score 51; DB 15; Length 163;
Best Local Similarity 50.0%; Pred. No. 5.9;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGTQPLILAQFMRVGGDELL 20
DB 22 AGLEPLLKARFGLAGGDL 41

RESULT 5

US-09-884-260A-11
Sequence 11, Application US/09884260A
Patent No. US20020098570A1
GENERAL INFORMATION:
APPLICANT: Alan Brash
APPLICANT: Nathalie Tijet
TITLE OF INVENTION: MUSKMELOX (CUCUMIS MELO) HYDROPEROXIDE
FILE REFERENCE: 06027.000202
CURRENT APPLICATION NUMBER: US/09/884,260A
CURRENT FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: 09/537,357
PRIOR FILING DATE: 2000-03-29
NUMBER OF SEQ ID NOS: 56
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 11
TYPE: PRT
LENGTH: 62
ORGANISM: Capsicum annuum (green pepper)
US-09-884-260A-11

Query Match 37.9%; Score 50; DB 9; Length 62;
Best Local Similarity 31.6%; Pred. No. 2.9;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPLI-----LAQFMRVGGDELLHFLW 25
DB 6 GYQPLVWKDPKVFDEPEKFMLEFRTKKGKELLNLYLFW 43

RESULT 6

US-10-425-115-327061
Sequence 327061, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 327061

```

; LENGTH: 70
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MWT4577_6134C.1.pep
US-10-425-115-327061

Query Match      37.9%; Score 50; DB 17; Length 70;
Best Local Similarity 47.1%; Pred. No. 3.4;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AGTQPLILAQPMRVGGD 17
Db 23 ASVQFVLIQTMVSGGE 39

RESULT 7
US-10-381-870-5
; Sequence 5, Application US/10381870
; Publication No. US20040088752A1
; GENERAL INFORMATION:
; APPLICANT: Howe, Gregg
; TITLE OF INVENTION: Divinyl Ether Synthase Gene, Protein, and Uses Thereof
; FILE REFERENCE: MSU-06815
; CURRENT APPLICATION NUMBER: US/10/381,870
; CURRENT FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: 60/238,415
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-10-381-870-5

Query Match      37.9%; Score 50; DB 15; Length 476;
Best Local Similarity 31.6%; Pred. No. 28;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

Qy 2 GTQPLI-----ILAQPMRVGGDELLHFLW 25
Db 385 GYQPLVMKDPKVFDEPEKFMLEFRTKKGKELLNLYFW 422

RESULT 8
US-09-884-260A-32
; Sequence 32, Application US/09884260A
; Patent No. US20020098570A1
; GENERAL INFORMATION:
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: MUSKXELON (CUCUMIS MELO) HYDROPEROXIDE
; FILE REFERENCE: 06027.0002U2
; CURRENT APPLICATION NUMBER: US/09/884,260A
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 09/537,357
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-09-884-260A-32

Query Match      37.9%; Score 50; DB 9; Length 480;
Best Local Similarity 31.6%; Pred. No. 28;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

Qy 2 GTQPLI-----LAQPMRVGGDELLHFLW 25

```

```

Db 389 GYQPLVMKDPKVFDEPEKFMLEFRTKKGKELLNLYFW 426

RESULT 9
US-10-042-991-25
; Sequence 25, Application US/10042991
; Publication No. US20020142407A1
; GENERAL INFORMATION:
; APPLICANT: Ian M. Whitehead
; APPLICANT: Alan Slusarenko
; APPLICANT: Duncan Gaskin
; APPLICANT: Alan Brash
; TITLE OF INVENTION: GUAVA (PSIDIUM GUAJAVA) 13-HYDROPEROXIDE
; TITLE OF INVENTION: LYASE AND USES THEREOF
; FILE REFERENCE: 06027.0001U3
; CURRENT APPLICATION NUMBER: US/10/042,991
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Capsicum annuum (green pepper)
US-10-042-991-25

Query Match      37.9%; Score 50; DB 13; Length 480;
Best Local Similarity 31.6%; Pred. No. 28;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

Qy 2 GTQPLI-----LAQPMRVGGDELLHFLW 25
Db 389 GYQPLVMKDPKVFDEPEKFMLEFRTKKGKELLNLYFW 426

RESULT 10
US-10-767-701-55335
; Sequence 55335, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 55335
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(103)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: 30161891.pep
US-10-767-701-55335

Query Match      36.4%; Score 48; DB 16; Length 103;
Best Local Similarity 55.0%; Pred. No. 11;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 6 LILAQPMRVGGDELLHFLW 25
Db 24 LILAIRRRKGSALLRILLW 43

RESULT 11
US-10-425-114-53219

```

/ APPLICANT: Burgess, Catherine E.
 / APPLICANT: Casman, Stacie J.
 / APPLICANT: Casman, William M.
 / APPLICANT: Gusev, Vladimir Y.
 / APPLICANT: J.L. Weizhen
 / APPLICANT: Lepley, Denise M.
 / APPLICANT: Liu, Xiaohong
 / APPLICANT: Mezick, Amanda J.
 / APPLICANT: Padigaru, Muralidhara
 / APPLICANT: Patturajan, Meera
 / APPLICANT: Rastelli, Luca
 / APPLICANT: Shen, Lei
 / APPLICANT: Shenoy, Suresh G.
 / APPLICANT: Shimkets, Richard A.
 / APPLICANT: Spaderna, Steven K.
 / APPLICANT: Spytek, Kimberly A.
 / APPLICANT: Szekeres, Edward S. Jr.
 / APPLICANT: Taupier, Raymond J. Jr.
 / APPLICANT: Tchernev, Veizar I.
 / APPLICANT: Zernusen, Bryan D.
 / APPLICANT: Voss, Edward Z.
 / TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

```

;
; CURRENT APPLICATION NUMBER: US/10/190,115
;
; CURRENT FILING DATE: 2003-02-10
;
; PRIOR APPLICATION NUMBER: 60/303,168
;

```

; PRIOR APPLICATION NUMBER: 60/368,996
 ;
 ; PRIOR FILING DATE: 2002-04-01
 ;
 ; PRIOR APPLICATION NUMBER: 60/386,816
 ;

```

, PRIOR FILING DATE: 2000-07-03
, PRIOR APPLICATION NUMBER: 60/215,856
, PRIOR FILING DATE: 2000-07-03
, PRIOR APPLICATION NUMBER: 60/215,902
, PRIOR FILING DATE: 2000-07-03
, PRIOR APPLICATION NUMBER: 60/216,585,
, PRIOR FILING DATE: 2000-07-07
, PRIOR APPLICATION NUMBER: 60/216,586
, PRIOR FILING DATE: 2001-07-07
, PRIOR APPLICATION NUMBER: 60/216,722
, PRIOR FILING DATE: 2000-07-07
, PRIOR APPLICATION NUMBER: 60/218,622
, PRIOR FILING DATE: 2000-07-17
, Remaining Prior Application data removed - See File Wrapper or PALM.
, NUMBER OF SEQ ID NOS: 136
, SOFTWARE: Curasequest version 0.1
, SEQ ID NO 28
, LENGTH: 2447
, TYPE: PRT
, ORGANISM: Homo sapiens
US-10-190-115-28

Query Match          36.4%; Score 48; DB 14; Length 2447;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;

```

QY 2 GTQPLILAQFMRVGG--DELLH 21

RESULT 14

Publication No. US20040014081A1
GENERAL INFORMATION:
APPLICANT: Alsbrook II, John P
APPLICANT: Spaderna, Stephen K
APPLICANT: Tchernev, Velizar
APPLICANT: Liu, Xiaohong
APPLICANT: Shenoy, Suresh

```
; APPLICANT: Spvtek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: NO. US20040014081A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 2447
; TYPE: PRT
; ORGANISM: Takifugu rubripes
US-10-369-072-28

Query Match      36.4%; Score 48; DB 15; Length 2447;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;
Matches 12; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY      2 GTQPLILAQFMVRVGG--DELLH 21
Db      1908 GTPLILARLAVEGVVEELIH 1929

RESULT 15
US-10-306-762-214
; Sequence 214, Application US/10306762
; Publication No. US20030187220A1
; GENERAL INFORMATION:
; APPLICANT: Park, Frances
; APPLICANT: Gajiwala, Ketan S.
; APPLICANT: Buchanan, Sean Grant
; APPLICANT: Sauder, J. Michael
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF A FLAVIN
; TITLE OF INVENTION: MONONUCLEOTIDE BINDING PROTEIN (FMNBP)
; FILE REFERENCE: 52498-20011.00
; CURRENT APPLICATION NUMBER: US/10/306,762
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US 60/334,132
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 242
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 214
```

```
; LENGTH: 315
; TYPE: PRT
; ORGANISM: D hafniense (23116726)
US-10-306-762-214

Query Match      35.6%; Score 47; DB 14; Length 315;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      2 GTQPLILAQFMVRVGGD 17
Db      62 GSEPALLAQAAARIGAD 77

Search completed: November 11, 2004, 07:41:45
Job time : 60.4093 secs
```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 41.9039 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-66
Perfect score: 132
Sequence: 1 AGTQPLLAQFMVGGDELLHFLW 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	38.6	951	2 Q6FK21	Q6fk21 candida gla
2	50	37.9	456	2 Q73YV9	Q73yv9 mycobacteri
3	50	37.9	456	2 AAS04161	Aas04161 mycobacte
4	50	37.9	476	2 Q9LLA9	Q9lla9 lycopersico
5	50	37.9	476	2 Q8ARH8	Q8arh8 lycopersico
6	50	37.9	479	2 Q8XGJ8	Q8xgj8 lycopersico
7	50	37.9	480	2 Q39443	Q39443 capsicum an
8	50	37.9	480	2 Q9ARH7	Q9arh7 capsicum an
9	50	37.9	496	2 Q93YF8	Q93yf8 nicotiana a
10	49	37.1	188	2 Q7UUN1	Q7uun1 rhodopirell
11	49	37.1	216	2 Q868N1	Q868n1 ixodes scap
12	49	37.1	573	2 Q88294	Q88294 pyrococcus
13	49	37.1	1078	2 Q27807	Q27807 tripeustes
14	49	37.1	1249	2 Q85152	Q85152 african swi
15	49	37.1	9376	2 Q85168	Q85168 pseudomonas
16	48	36.4	76	2 Q8CYZ9	Q8cyz9 streptococc
17	48	36.4	236	2 Q95SS8	Q95ss8 drosophila
18	48	36.4	382	2 Q79V73	Q79v73 esalmonella
19	48	36.4	382	2 Q2Z435	Q2z435 plasmid col
20	48	36.4	382	2 Q7DJM3	Q7djm3 plasmid r64
21	48	36.4	480	2 Q93X18	Q93x18 solanum tub
22	48	36.4	499	2 Q76LM3	Q76lm3 citrus jamb
23	48	36.4	499	2 Q84V86	Q84v86 citrus sine
24	48	36.4	499	2 BAC55161	Bac55161 citrus ja
25	48	36.4	630	2 Q70SU0	Q70su0 suberites d
26	48	36.4	630	2 CAD79441	Cad79441 suberites
27	48	36.4	630	2 CAD66418	Cad66418 suberites
28	48	36.4	819	2 Q6GLQ4	Q6glq4 xenopus lae
29	48	36.4	1036	1 SNO DROME	P91682 drosophila
30	48	36.4	1036	2 Q86PA9	Q86pa9 drosophila
31	48	36.4	2447	2 O13149	O13149 figu rubrip

32	47	35.6	235	2 Q35542	Q35542 nereis dive
33	47	35.6	371	2 Q8HT81	Q8ht81 selaginella
34	47	35.6	409	2 Q6BML9	Q6bml9 debaryomyce
35	47	35.6	480	2 Q9RV87	Q9rv87 deinococcus
36	47	35.6	569	2 Q8AVK1	Q8avk1 lactobacill
37	47	35.6	625	2 Q8KUK1	Q8kuk1 vibrio chol
38	47	35.6	740	2 Q8PG03	Q8pg03 xanthomonas
39	47	35.6	874	1 SYL_BRAJA	Q89wq1 bradyrhizob
40	47	35.6	878	2 Q7SIE4	Q7sie4 thermus the
41	47	35.6	878	2 Q7ZGM3	Q7zgm3 thermus the
42	47	35.6	878	2 AAS82167	Aas82167 thermus t
43	46.5	35.2	457	2 Q896Q1	Q896q1 clostridium
44	46	34.8	232	2 Q82XM2	Q82xm2 pyrobaculum
45	46	34.8	294	2 Q7CWZ9	Q7cwz9 agrobacteri

ALIGNMENTS

RESULT 1					
Q6FK21	PRELIMINARY;	PRT;	951	AA.	
ID	Q6FK21				
AC	Q6FK21;				
DT	05-JUL-2004 (Tremblrel. 27, Created)				
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)				
DT	05-JUL-2004 (Tremblrel. 27, Last annotation update)				
DE	Similar to sp P38249 Saccharomyces cerevisiae YER079c RPL1.				
GN	ORFNames=CAGL0M018049;				
OS	Candida glabrata (yeast) (Torulopsis glabrata).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; mitosporic Saccharomycetales; Candida.				
OX	NCBI_TaxID=5478;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CBS138;				
RG	GENOLEVURES;				
RA	Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,				
RA	Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,				
RA	Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,				
RA	Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,				
RA	Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,				
RA	Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,				
RA	Kantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,				
RA	Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,				
RA	Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,				
RA	Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,				
RA	Swennene D., Tekala F., Wesolowski-Louhel M., Westhof E., Wirth B.,				
RA	Zenou-Meyer M., Zivanovic I., Biotin-Fukuhara M., Thierry A.,				
RA	Bouchier C., Caudron B., Scarcelli C., Gaillardin C., Weissenbach J.,				
RA	Wincker P., Souciet J.L.;				
RT	"Genome evolution in yeasts."				
RL	Nature 430:35-44 (2004).				
DR	EMBL; CR380959; CAG62399.1;				
DR	InterPro; IPR010982; Lambda_like_DNA.				
DR	InterPro; IPR000717; PCI.				
DR	SMART; SM00088; PINT; I.				
SQ	SEQUENCE 951 AA; 110156 MW; DC6B47AEBB351A5A CRC64;				
Query Match 38.6%; Score 51; DB 2; Length 951;					
Best Local Similarity 50.0%; Pred. No. 56;					
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;					
QY 8 LAQFMVGGDELLHFLW 25					
DB 286 LVQIFLVGGDIHTLW 303					
RESULT 2					
Q73YV9	PRELIMINARY;	PRT;	456	AA.	
ID	Q73YV9				
AC	Q73YV9;				
DT	05-JUL-2004 (Tremblrel. 27, Created)				
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)				

```

DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE Lpd.
GN Name=lpd; OrderedLocusNames=MAP1844c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: FAD (By similarity);
CC -!- SIMILARITY: Belongs to the class-I pyridine nucleotide-disulfide
CC oxidoreductase family.
DR EMBL; AE017233; AAS04161.1; -.
DR InterPro; IPR001327; FAD pyr redox.
DR InterPro; IPR000815; Hg reductase.
DR InterPro; IPR001100; Pyr redox.
DR InterPro; IPR004099; Pyr_redox_dim.
DR Pfam; PF00070; Pyr_redox; 1.
DR Pfam; PF02852; Pyr_redox_dim; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR0945; HGRDTASEI.
DR PRODOM; PD000139; FAD pyr redox; 1.
DR PROSITE; PS00076; PYRIDINE_REDOX_1; 1.
DR Complete proteome; FAD; Flavoprotein; Oxidoreductase;
KW Redox-active center.
KW Redox-active center.
SQ SEQUENCE 456 AA; 48736 MW; B36BB27F78C2451E CRC64;

Query Match 37.9%; Score 50; DB 2; Length 456;
Best Local Similarity 41.7%; Pred. No. 38;
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 ACTQPLILAQPMRVGCGDELLHFL 24
DB 396 ADTBILGAAILGVGGDEAIRGIL 419

RESULT 3
AAS04161 PRELIMINARY; PRT; 456 AA.
ID AAS04161
AC AAS04161;
DT 02-MAR-2004 (TREMELrel. 27, Created)
DT 02-MAR-2004 (TREMELrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMELrel. 27, Last annotation update)
DE Lpd.
GN LPD OR MAP1844c.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017233; AAS04161.1; -.
SQ SEQUENCE 456 AA; 48736 MW; B36BB27F78C2451E CRC64;

Query Match 37.9%; Score 50; DB 2; Length 456;
Best Local Similarity 41.7%; Pred. No. 38;
Matches 10; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 ACTQPLILAQPMRVGCGDELLHFL 24
DB 396 ADTBILGAAILGVGGDEAIRGIL 419

RESULT 4
Q9LLA9 PRELIMINARY; PRT; 476 AA.
ID Q9LLA9
AC Q9LLA9;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Fatty acid hydroperoxide lyase.
GN Name=HPL;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20317213; PubMed=10859201;
RA Howe G.A., Lee G.I., Itoh A., Li L., DeRoche A.E.;
RT "Cytochrome P450-dependent metabolism of oxylipins in tomato. Cloning
RT and expression of allene oxide synthase and fatty acid hydroperoxide
RT lyase.";
RL Plant Physiol. 123:711-724(2000).
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AF230372; AAF67142.1; -.
DR FIR; JC7304; JC7304.
DR GO; GO:0016829; F-lyase activity; IEA.
DR GO; GO:0004497; F-monooxygenase activity; IEA.
DR GO; GO:0006118; P-electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00465; EP450IV.
KW Heme; Lyase.
KW Heme; Lyase.
SQ SEQUENCE 476 AA; 53542 MW; AB0EA80F0C97C4EF CRC64;

Query Match 37.9%; Score 50; DB 2; Length 476;
Best Local Similarity 31.6%; Pred. No. 39;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPL-----ILAQPMRVGCGDELLHFLW 25
DB 385 GYQPLVMKDPKVFDEPKFVLEFRTKRGKELLNLYFW 422

RESULT 5
Q9ARH8 PRELIMINARY; PRT; 476 AA.
ID Q9ARH8
AC Q9ARH8;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Fatty acid hydroperoxide lyase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Richard S., Atwal A.S.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AY028373; AAK27265.1; -.
DR FIR; JC7304; JC7304.
DR GO; GO:0016829; F-lyase activity; IEA.
DR GO; GO:0004497; F-monooxygenase activity; IEA.
DR GO; GO:0006118; P-electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00465; EP450IV.
KW Heme; Lyase.
KW Heme; Lyase.
SQ SEQUENCE 476 AA; 53480 MW; 4A56DDB8131FB1C1 CRC64;

```


Query Match 37.9%; Score 50; DB 2; Length 476;
Best Local Similarity 31.6%; Pred. No. 39;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPLI-----ILAQFMRVGGDELLHFLW 25
| | | | | : : : : :
Db 385 GYQPLVMKDPKVFDEPEKFLERFTKEKGKELLNLFW 422

RESULT 6
Q9XG18 PRELIMINARY; PRT; 479 AA.
ID Q9XG18
AC Q9XG18
DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hydroperoxide lyase (Fragment).
GN Name=hpl;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Sheldrick B.;
RL Genetics, Guelph, ON, Canada.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AJ239065; CAB33022.1; -.
DR FIR; JC7304; JC7304.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00465; EP450IV.
KW Heme; Lyase.
FT NON TER
SQ SEQUENCE 479 AA; 54012 MW; ED6C06DA1DE81953 CRC64;

Query Match 37.9%; Score 50; DB 2; Length 479;
Best Local Similarity 31.6%; Pred. No. 40;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPLI-----ILAQFMRVGGDELLHFLW 25
| | | | | : : : : :
Db 388 GYQPLVMKDPKVFDEPEKFLERFTKEKGKELLNLFW 425

RESULT 7
Q39443 PRELIMINARY; PRT; 480 AA.
ID Q39443
AC Q39443
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Fatty acid hydroperoxide lyase.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fruits;
RA Matsui K., Shibutani M., Kajiwara T., Hase T.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; U51674; AA97465.1; -.
DR GO; GO:0016829; F:lyase activity; IEA.

DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00465; EP450IV.
KW Heme; Lyase.
SQ SEQUENCE 480 AA; 54056 MW; 84A2F646A55D46C9 CRC64;

Query Match 37.9%; Score 50; DB 2; Length 480;
Best Local Similarity 31.6%; Pred. No. 40;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPLI-----LAQFMRVGGDELLHFLW 25
| | | | | : : : : :
Db 389 GYQPLVMKDPKVFDEPEKFLERFTKEKGKELLNLFW 426

RESULT 8
Q9ARH7 PRELIMINARY; PRT; 480 AA.
ID Q9ARH7
AC Q9ARH7
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Fatty acid hydroperoxide lyase.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Richard S., Atwal A.S.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AY028374; AAK27266.1; -.
DR FIR; S74228; S74228.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00465; EP450IV.
KW Heme; Lyase.
SQ SEQUENCE 480 AA; 54125 MW; EFC8646A55D4798 CRC64;

Query Match 37.9%; Score 50; DB 2; Length 480;
Best Local Similarity 31.6%; Pred. No. 40;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTQPLI-----LAQFMRVGGDELLHFLW 25
| | | | | : : : : :
Db 389 GYQPLVMKDPKVFDEPEKFLERFTKEKGKELLNLFW 426

RESULT 9
Q93YF8 PRELIMINARY; PRT; 496 AA.
ID Q93YF8
AC Q93YF8
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Hydroperoxide lyase.
GN Name=hpl;
OS Nicotiana attenuata.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=49451;
RN [1]

```

RP SEQUENCE FROM N.A.
RA Keinaenen M.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ziegler J.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AJ414400; CAC91565.1; -.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; F:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR002403; EP450IV.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PRO0465; EP450IV.
KW Heme; Lyase.
SQ SEQUENCE 496 AA; 55534 MW; 5061E4B0919B0921 CRC64;

Query Match 37.9%; Score 50; DB 2; Length 496;
Best Local Similarity 31.6%; Pred. No. 41;
Matches 12; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 GTGPI-----IIAQFMRVGGDELLHFLW 25
   |||
   |||
Db 405 GYQPLVMDPKVDDPDKFVLERFTKKGKELLNLFW 442

RESULT 10
Q7UUN1 PRELIMINARY; PRT; 188 AA.
AC Q7UUN1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=RB3194;
OS Rhodopirellula baltica
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Firellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Firellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
RW EMBL; BX294138; CAD73048.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 188 AA; 21276 MW; 34C149E43989435C CRC64;

Query Match 37.1%; Score 49; DB 2; Length 188;
Best Local Similarity 47.4%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 7 ILAQFMRVGGDELLHFLW 25
   :|||
   :|||
Db 140 LURQFFAFGGDDRQQFSLW 158

RESULT 11
Q868N1 PRELIMINARY; PRT; 216 AA.
AC Q868N1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Is4.
OS Ixodes scapularis (Black-legged tick) (Deer tick).

```

```

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Ixodes.
OX NCBI_TaxID=6945;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22480955; PubMed=12593519;
RA Packila M., Guilfoile P.G.;
RT "Mating, male Ixodes scapularis express several genes including those
RT with sequence similarity to immunoglobulin-binding proteins and
RT metalloproteases."
RL Exp. Appl. Acarol. 27:151-160(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Packila M., Guilfoile P.G.; EMBL/GenBank/DBJ databases.
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY234848; AAO85921.1; -.
SQ SEQUENCE 216 AA; 23990 MW; C9C97B9B331FE0AE CRC64;

Query Match 37.1%; Score 49; DB 2; Length 216;
Best Local Similarity 55.6%; Pred. No. 25;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 8 LAQFMRVGGDELLHFLW 25
   :|||
   :|||
Db 198 LSRWNRHGGHELEFLW 215

RESULT 12
O58294 PRELIMINARY; PRT; 573 AA.
AC O58294;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein PH0559.
GN OrderedLocNames=PH0559;
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
RW EMBL; AP000002; BAA29648.1; -.
DR EMBL; C71170; C71170.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR InterPro; IPR001952; Bact_sec_systII.
DR Pfam; PF00482; GSP11_F; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 573 AA; 64304 MW; 91C98E6BEF45CC36 CRC64;

Query Match 37.1%; Score 49; DB 2; Length 573;
Best Local Similarity 47.4%; Pred. No. 69;
Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 6 LIIAQFMRVGGDELLHFLW 24
   :|||
   :|||
Db 553 VLMSKFMVGGGSLILFLW 571

RESULT 13
Q27807

```

ID Q27807 PRELIMINARY; PRT; 1078 AA.
AC Q27807;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Dynein heavy chain isoform 5A (EC 3.6.1.3) (Fragment).
GN Name=DYHSA;
OS Tripneustes gratilla (Hawaiian sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Tripneustes.
OX NCBI_TaxID=7673;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=94243035; PubMed=8186465;
RA Gibbons B.H., Asai D.J., Tang W.J., Hays T.S., Gibbons I.R.;
RT "Phylogeny and expression of axonemal and cytoplasmic dynein genes in
RT sea urchins".
RL MOL. Biol. Cell 5:57-70(1994).
DR EMBL; U03977; AAA63591.1; -.
DR PIR; T30879; T30879.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042624; F:ATPase activity, uncoupled; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR SMART; SM00382; AAA; 3.
KW ATP-binding; Hydrolase.
FT NON_TER 1
FT NON_TER 1078
FT SEQUENCE 1078 AA; 121418 MW; 049AE4EA66316329 CRC64;
SQ

Query Match 37.1%; Score 49; DB 2; Length 1078;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 3 TOPLIQAQFMVGGDE 18
|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
Db 939 TKPIFGDFMKVGAE 954

RESULT 14
Q65152
ID Q65152 PRELIMINARY; PRT; 1249 AA.
AC Q65152;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE PM1249L.
GN Name=M1249L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=90223993; PubMed=2327074;
RA Lopez-Otin C., Freije J.M.P., Parra F., Mendez E., Vinuela E.;
RT "Mapping and sequence of the gene coding for protein p72, the major
RT capsid protein of African swine fever virus.";
RL Virology 175:477-484(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=90357780; PubMed=2389555;
RA Blasco R., Lopez-Otin C., Munoz M., Bockamp E.-O.O., Simon-Mateo C.,
RA Vinuela E.;
RT "Sequence and evolutionary relationships of African swine fever virus
RT thymidine kinase.";
RL Virology 178:301-304(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX

RC STRAIN=BA71V;
RX MEDLINE=90219204; PubMed=2325202;
RA Almendral J.M., Almazan F., Blasco R., Vinuela E.;
RT "Multigene families in African swine fever virus: family 110.";
RL J. Virol. 64:2064-2072(1990).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=90219205; PubMed=2325203;
RA Gonzalez A., Calvo V., Almazan F., Almendral J.M., Ramirez J.C.,
RA De la Vega I., Blasco R., Vinuela E.;
RT "Multigene families in African swine fever virus: Family 360.";
RL J. Virol. 64:2073-2081(1990).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=91134988; PubMed=1994575;
RA Camacho A., Vinuela E.;
RT "Protein p22 of African swine fever virus: an early structural protein
RT that is incorporated into the membrane of infected cells.";
RL Virology 181:251-257(1991).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=92287485; PubMed=1309282;
RA Rodriguez J.M., Salas M.L., Vinuela E.;
RT "Genes homologous to ubiquitin-conjugating proteins and eukaryotic
RT transcription factor SII in African swine fever virus.";
RL Virology 186:40-52(1992).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93174976; PubMed=8438592;
RA Yanez R.J., Vinuela E.;
RT "African swine fever virus encodes a DNA ligase.";
RL Virology 193:531-536(1993).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=92263807; PubMed=1316688;
RA Garcia-Beato E., Freije J.M.P., Lopez-Otin C., Blasco R., Vinuela E.,
RA Salas M.L.;
RT "A gene homologous to topoisomerase II in african swine fever virus.";
RL Virology 188:938-947(1992).
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=92260660; PubMed=1583732;
RA Alami A., Angulo A., Lopez-Otin C., Munoz M., Freije J.M.P.,
RA Carrascosa A.L., Vinuela E.;
RT "Amino acid sequence and structural properties of protein p12, an
RT African swine fever virus attachment protein.";
RL J. Virol. 66:3860-3868(1992).
RN [10]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=94091056; PubMed=8266720;
RA Freije J.M., Lain S., Vinuela E., Lopez-Otin C.;
RT "Nucleotide sequence of a nucleoside triphosphate phosphohydrolase
RT gene from African swine fever virus.";
RL Virus Res. 30:63-72(1993).
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93346971; PubMed=8393914;
RA Yanez R.J., Rodriguez J.M., Rodriguez J.F., Salas M.L., Vinuela E.;
RT "African swine fever virus thymidylate kinase gene: sequence and
RT transcriptional mapping.";
RL J. Gen. Virol. 74:1633-1638(1993).
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93281390; PubMed=8506138;

```

RX MEDLINE=94187118; PubMed=8139051;
RA Rodriguez J.M., Yanez R.J., Pan R., Rodriguez J.F., Salas M.L.,
RA Vinuela E.;
RT "Multigene families in African swine fever virus: family 505.";
RT J. Virol. 68:2746-2751(1994).
RN [22]
RP SEQUENCE FROM N.A.
RC
RC STRAIN=BA71V;
RC MEDLINE=94233765; PubMed=8178480;
EX De la Vega I., Gonzalez A., Blasco R., Calvo V., Vinuela E.;
RT "Nucleotide sequence and variability of the inverted terminal
RT repetitions of African swine fever virus DNA.";
RT Virology 201:152-156(1994).
RN [23]
RP SEQUENCE FROM N.A.
RC
RC STRAIN=BA71V;
RC MEDLINE=21820291; PubMed=11831707;
EX Yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C.,
RA Rodriguez J.F., Vinuela E.;
RT "Analysis of the complete nucleotide sequence of African swine fever
RT virus.";
RT Virology 208:249-278(1995).
RN [24]
RP SEQUENCE FROM N.A.
RC
RC STRAIN=BA71V;
RC
RC Query Match 37.1%; Score 49; DB 2; Length 1249;
RC Best Local Similarity 45.0%; Pred. No. 1.6e-02;
RC Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps
Qy 4 QPILIAQMRVGGDELLHFL 23
Db 488 QGLIYSYFCIGGELAHFI 507

RESULT 15
OR5168 PRELIMINARY; PRT; 9376 AA.
ID 085168;
AC 085168;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Syringomycin synthetase.
DE Name-syS;
OS Pseudomonas syringae (pv. syringae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=321;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99047670; PubMed=9830033;
RA Guenzi E., Galli G., Grgurina I., Gross D.C., Grandi G.;
RT "Characterization of the syringomycin synthetase gene cluster. A link
RT between prokaryotic and eukaryotic peptide synthetases.";
RT J. Biol. Chem. 273:32857-32863(1998).
RC J-1. SIMILARITY: 273 to the ATP-dependent AMP-binding enzyme
CC family
CC ENML; AF047828; AAC80285.1; -.
DR PIR; T14593; T14593.
DR HSP; P14687; IAMU.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0016788; F:hydrolase activity, acting on ester bonds; IEA.
DR GO; GO:0003058; P:biosynthesis; IEA.
DR InterPro; IPR010071; AA_adenyl_dom.
DR InterPro; IPR009881; ACP_like.
DR InterPro; IPR000873; AMP_bind.
DR InterPro; IPR001242; Condensatn.
DR InterPro; IPR006162; Ppantne.S.
DR InterPro; IPR006163; Pp_bind.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00501; AMP-binding; 8.
DR Pfam; PF00668; Condensation; 9.

```

DR Pfam; PF00550; PP-binding; 9.
 DR Pfam; PF00975; Thioesterase; 1.
 DR PRINTS; PR00154; AMPBINDING.
 DR TIGRFAMs; TIGR01733; AA-adenyl-dcm; 8.
 DR PROSITE; PS00075; ACP DOMAIN; 9.
 DR PROSITE; PS00455; AMP BINDING; 8.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 5.
 KW Phosphopantetheine.
 SQ SEQUENCE 9376 AA; 1029843 MW; F770C08975EF9CE5 CRC64;
 Query Match 37.1%; Score 49; DB 2; Length 9376;
 Best Local Similarity 43.5%; Pred.No. 1.3e+03;
 Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 Qy 1 AGTQPLILAQFMRYVGGDELLHFL 23
 Db 8155 SGHKLVIIPQLRANGSELDFL 8177

Search completed: November 10, 2004, 15:53:26
 Job time : 43.9808 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 14.8043 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-67
Perfect score: 158
Sequence: 1 MDTIKGFDLITNFQVVADALNISLLENPLATA 32

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PGTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfilee1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.5	34.5	523	4	US-09-328-352-6394
2	48.5	30.7	221	4	US-09-248-796A-20200
3	48	30.4	425	4	US-09-252-991A-23329
4	47	29.7	271	4	US-09-266-965-141
5	47	29.7	309	4	US-09-540-236-2848
6	47	29.7	692	4	US-09-248-796A-17390
7	46.5	29.4	523	4	US-09-328-352-6395
8	46	29.1	63	4	US-09-489-039A-14297
9	46	29.1	451	4	US-09-134-000C-3849
10	46	29.1	744	4	US-09-248-796A-18090
11	45	28.5	125	4	US-09-270-767-34186
12	45	28.5	125	4	US-09-270-767-49403
13	45	28.5	453	2	US-08-484-126-3
14	45	28.5	453	4	US-09-374-909-3
15	45	28.5	475	4	US-09-248-796A-15218
16	45	28.5	575	1	US-08-403-866-7
17	45	28.5	645	4	US-09-315-127-8
18	45	28.5	645	4	US-09-315-127-9
19	44.5	28.2	510	4	US-09-710-279-2860
20	44.5	28.2	513	3	US-09-134-001C-4490
21	44	27.8	113	4	US-09-710-279-2338
22	44	27.8	114	4	US-09-513-999C-5753
23	44	27.8	134	4	US-09-621-976-6645
24	44	27.8	141	4	US-09-252-991A-21958
25	44	27.8	172	4	US-09-248-796A-16569
26	44	27.8	203	4	US-09-248-796A-21030
27	44	27.8	246	4	US-09-540-236-2695

28	44	27.8	259	4	US-09-522-714-18	Sequence 18, Appli
29	44	27.8	260	4	US-09-710-279-224	Sequence 224, App
30	44	27.8	271	4	US-09-688-019-4	Sequence 4, Appli
31	44	27.8	283	4	US-09-602-787A-588	Sequence 588, App
32	44	27.8	302	3	US-09-134-001C-3171	Sequence 3171, Ap
33	44	27.8	314	4	US-09-270-767-42409	Sequence 42409, A
34	44	27.8	465	4	US-09-328-352-4543	Sequence 4543, Ap
35	44	27.8	508	4	US-09-252-991A-32765	Sequence 32765, A
36	44	27.8	534	4	US-09-248-796A-15081	Sequence 15081, A
37	44	27.8	859	3	US-09-149-934-1	Sequence 1, Appli
38	44	27.8	1248	4	US-10-042-810-2	Sequence 2, Appli
39	44	27.8	1278	4	US-10-042-810-4	Sequence 4, Appli
40	44	27.8	1338	4	US-09-631-603-2	Sequence 2, Appli
41	44	27.8	2629	2	US-08-751-189-4	Sequence 4, Appli
42	44	27.8	2629	2	US-09-060-836-4	Sequence 4, Appli
43	44	27.8	2629	3	US-09-184-445-4	Sequence 4, Appli
44	43.5	27.5	196	4	US-09-248-796A-17562	Sequence 17562, A
45	43.5	27.5	302	4	US-09-662-746A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-328-352-6394
; Sequence 6394, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6394
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6394

Query Match 34.5%; Score 54.5; DB 4; Length 523;
Best Local Similarity 34.9%; Pred. No. 3.1;
Matches 15; Conservative 5; Mismatches 12; Indels 11; Gaps 1;
QY 1 MDTIKGFDLITNFQV-----VADALNISLLENPLATA 32
DB 110 LNQIKGLNKANFDVFSLSCHNCDFVVQALNLIAYNPNTTA 152

RESULT 2
US-09-248-796A-20200
; Sequence 20200, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20200
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20200

Query Match 30.7%; Score 48.5; DB 4; Length 221;


```

; SEQ ID NO 6395
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6395

Query Match          29.4%; Score 46.5; DB 4; Length 523;
Best Local Similarity 32.6%; Pred No. 58;
Matches 14; Conservative 5; Mismatches 13; Indels 11; Gaps 1;

QY 1 MDTIKGFDLITNFQV-----VADALNISLLPLNPLATA 32
   : : : : : : : : : : : : : : : : : : : :
Db 110 LNQIKGLNLANKFDFVSLSCHNCPCDVVQALNLIAIYNSNTTA 152

RESULT 8
US-09-489-039A-14297
; Sequence 14297, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 14297
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-14297

Query Match          29.1%; Score 46; DB 4; Length 63;
Best Local Similarity 56.5%; Pred. NO. 4.4;
Matches 13; Conservative 3; Mismatches 5; Indels 2; Gaps 2;

QY 6 GFDLITNFQVADALNISLLP 28
   : : : : : : : : : : : : : : : :
Db 9 GFGVTFNQVLL-AL-ASLIPT 29

RESULT 9
US-09-134-000C-3849
; Sequence 3849, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3849
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3849

Query Match          29.1%; Score 46; DB 4; Length 451;
Best Local Similarity 64.3%; Pred No. 58;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 ITNFQWADALNIS 23
   : : : : : : : : : :
Db 114 LNNFQKIADALTIS 127

```

```
; SEQ ID NO 49403
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49403

Query Match      28.5%; Score 45; DB 4; Length 125;
Best Local Similarity 41.7%; Pred. No. 16;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      2 DTIKGFDLITNQVQVADALNISLL 25
Db      26 NTVNARNLINHFSVVRIPIINIKLM 49

RESULT 13
US-08-484-126-3
; Sequence 3, Application US/08484126
; Patent No. 5985655
; GENERAL INFORMATION:
; APPLICANT: Anderson, W. French
; APPLICANT: Baltrucki, Leon F.
; APPLICANT: Mason, James M.
; TITLE OF INVENTION: Targetable Vector Particles
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,126
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,347
; FILING DATE: 20-OCT-1994
; APPLICATION NUMBER: 08/973,307
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lillie, Raymond J.
; REGISTRATION NUMBER: 31,778
; REFERENCE/DOCKET NUMBER: 271010-281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: xenotropic gp70 protein
; SEQUENCE DESCRIPTION: 271010-281

US-08-484-126-3
Query Match      28.5%; Score 45; DB 2; Length 453;
Best Local Similarity 33.3%; Pred. No. 84;
Matches 10; Conservative 5; Mismatches 11; Indels 4; Gaps 1;

QY      6 GFDLITNQVQVADALNIS----LLPNPLAT 31
Db      26 NTVNARNLINHFSVVRIPIINIKLM 49

RESULT 14
US-09-374-909-3
; Sequence 3, Application US/09374909
; Patent No. 6503501
; GENERAL INFORMATION:
; APPLICANT: Anderson, W. French
; APPLICANT: Baltrucki, Leon F.
; APPLICANT: Mason, James M.
; TITLE OF INVENTION: Targetable Vector Particles
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/374,909
; FILING DATE: 13-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/484,126
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/973,307
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lillie, Raymond J.
; REGISTRATION NUMBER: 31,778
; REFERENCE/DOCKET NUMBER: 271010-281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: xenotropic gp70 protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-374-909-3
Query Match      28.5%; Score 45; DB 4; Length 453;
Best Local Similarity 33.3%; Pred. No. 84;
Matches 10; Conservative 5; Mismatches 11; Indels 4; Gaps 1;

QY      6 GFDLITNQVQVADALNIS----LLPNPLAT 31
Db      225 GADPVTFRSLTRQVLNVGPRVPIGNPVIT 254

RESULT 15
US-09-248-796A-15218
; Sequence 15218, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
```

```

; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15218
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15218

Query Match      28.5%; Score 45; DB 4; Length 475;
Best Local Similarity 50.0%; Pred. No. 89;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      10 ITNFQVVADALNISILPN 27
Db      69 LTCFGTLADALNVRLKEH 86

Search completed: November 10, 2004, 15:57:20
Job time : 14.8043 secs

```

This Page Blank (up10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2004, 02:43:24 ; Search time 77.3238 Seconds
(without alignments)
146.426 Million cell updates/sec

Title: US-10-092-750-67

Perfected score: 158

Sequence: 1 MTIKGFDLITNFQVVDALNISLLPNPLATA 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	158	100.0	32	14	US-10-092-750-67
2	51.5	32.6	254	15	US-10-424-599-156639
3	51	32.3	171	17	US-10-425-115-367776
4	51	32.3	1097	15	US-10-424-599-189678
5	50	31.6	121	16	US-10-437-963-122995
6	50	31.6	426	17	US-10-425-115-243733
7	50	31.6	435	15	US-10-425-114-45323
8	49	31.0	87	17	US-10-425-115-316100
9	49	31.0	552	16	US-10-437-963-189437
10	48.5	30.7	207	16	US-10-767-701-42726
11	48.5	30.7	260	17	US-10-425-115-262020
12	48.5	30.7	1800	15	US-10-363-929-133
13	48.5	30.7	1800	15	US-10-367-502-242

Sequence 355295,
Sequence 172368,
Sequence 131609,
Sequence 183864,
Sequence 139, App
Sequence 132842,
Sequence 456, App
Sequence 2848, App
Sequence 141, App
Sequence 141, App
Sequence 13463, A
Sequence 176386,
Sequence 131113,
Sequence 131653,
Sequence 132618,
Sequence 131312,
Sequence 115293,
Sequence 180919,
Sequence 152989,
Sequence 241734,
Sequence 4754, Ap
Sequence 2412, Ap
Sequence 42, Appl
Sequence 119426,
Sequence 129864,
Sequence 100.0%; Score 158; DB 14;
Best Local Similarity 100.0%; Pred No. 7,3e-17;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTIKGFDLITNFQVVDALNISLLPNPLATA 32
Db 1 MTIKGFDLITNFQVVDALNISLLPNPLATA 32
RESULT 2
US-10-424-599-156639
; Sequence 156639, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

ALIGNMENTS

RESULT 1
US-10-092-750-67
; Sequence 67, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/05002
; CURRENT APPLICATION NUMBER: US/10/092,750
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-67

Query Match 100.0%; Score 158; DB 14;
Best Local Similarity 100.0%; Pred No. 7,3e-17;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTIKGFDLITNFQVVDALNISLLPNPLATA 32
Db 1 MTIKGFDLITNFQVVDALNISLLPNPLATA 32
RESULT 2
US-10-424-599-156639
; Sequence 156639, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

```

; NUMBER OF SEQ ID NOS: 285684
;
; SEQ ID NO 189678
;
; LENGTH: 1097
;
; TYPE: PRT
;
; ORGANISM: Glycine max
;
; FEATURE:
;
; NAME/KEY: unsure
;
; LOCATION: (1)..(1097)
;
; OTHER INFORMATION: unsure at all Xaa locations
;
; FEATURE:
;
; OTHER INFORMATION: Clone ID: PAT_MRT3847_142295C.1.pep
;
; US-10-434-599-189678

```

Query Match	32.3%	Score 51;	DB 15;	Length 1097;
Best Local Similarity	46.2%	Pred. No. 1.5e+02;		
Matches 12:	Conservative	3;	Mismatches 11;	Indels

RESULT 5
 US-10-437-963-122995
 ; Sequence 122995, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

```

; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 122995
; LENGTH: 121
; TYPE: PAT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_25872C.1.pep
US/10-437-963-122995

```

Query Match 31.6%; Score 50; DB 16; Length 121;
Best Local Similarity 64.7%; Pred. No. 14;
Matches 11: Conservative 3; Mismatches 3; Indels

QY	14 QVADALNISLLPNPLA 30 : : :
b	32 RWADAALLSLSPSPLA 48

```

RESULT 6
US-10-425-115-243733
; Sequence 243733, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION: Thomas J.
; APPLICANT: La Rosa, David K.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/435,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

```

```

; SEQ ID NO 243733
;
; LENGTH: 426
; TYPE: prt
; ORGANISM: Zea mays
;
; FEATURE:
;
;   OTHER INFORMATION: Clone ID: MR*4577_153857C.1.bep
US-10-425-115-243733

```

Query Match 31.6%; Score 50; DB 17; Length 426;
Best Local Similarity 43.5%; Pred. NO. 64;
Matches 10; Conservative 7; Mismatches 6; Indels

QY	8	DLITNFQVADALNISILPNPLA	30
	:	:	:
	:	:	:
	:	:	:
	:	:	:
	:	:	:
D'b	304	ELVAGQKIVSEALKISLLSDPRA	326

RESULT 7

```

US-10-425-114-45323
; Sequence 45323, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 45323
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17201A01_FLI pep
US-10-425-114-45323

```

Query Match 31.6%; Score 50; DB 15; Length 435;
Best Local Similarity 43.5%; Pred. No. 66;
Matches 10; Conservative 7; Mismatches 6; Indels

Qy 8 DLITNFQVADALNISLLPNPLA 30
:|: :|: :|: :|: :|: :|:
Db 313 ELVAGQKIVSEALKISLLSDPRA 335

RESIST.

```

RESULTS
US-10-425-115-316100
; Sequence 316100, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 316100
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRR4577 51354C.1.pep
US-10-425-115-316100

```

Query Match 31.0%; Score 49; DB 17; Length 87;
Best Local Similarity 54.5%; Pred. No. 13;
Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 11 TNFQVADALNISLLPNPLATA 32
 | | | | |
Dd 1 TFIQYADLVISMLOHPLATA 22

RESULT 9

```

RES001.1
US-10-437-963-189437
; Sequence 189437, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbatuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 189437
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85948C.1.pep
US-10-437-963-189437

```

Query Match 31.0%; Score 49; DB 16; Length 552;
Best Local Similarity 42.9%; Pred. No. 1.3e+02;
Matches 9; Conservative 5; Mismatches 7; Indels

QY	1	MDTIKGFDLITNFQVADALN	21
		: : : : : : : : :	
Db	68	LDVIRGMDWLTMFKGVIDCVN	88

RESULT 10

```

US-10-767-701-42726
; Sequence 42726, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 42726
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C118763_1.pep
; US-10-767-701-42726

```

Query Match 30.7%; Score 48.5; DB 16; Length 207;
Best Local Similarity 40.0%; Pred. No. 45;
Matches 12; Conservative 3; Mismatches 2; Indels 13; Gaps 1;

3 TIKGFDLITNFQVADALNISLLPNPLATA 32

```

Db      156 TVKGPDL-----:|::|| |::|  

RESULT 11  

US-10-425-115-262020  

; Sequence 262020, Application US/10425115  

; Publication No. US20040211472A1  

; GENERAL INFORMATION:  

; APPLICANT: La Rosa, Thomas J.  

; APPLICANT: Kovalic, David K.  

; APPLICANT: Zhou, Yihua  

; APPLICANT: Cao, Yongwei  

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  

; FILE REFERENCE: Plants  

; CURRENT APPLICATION NUMBER: B  

; CURRENT FILING DATE: 2003-04-28  

; NUMBER OF SEQ ID NOS: 369326  

; SEQ ID NO 262020  

; LENGTH: 260  

; TYPE: PRT  

; ORGANISM: Zea mays  

; FEATURE:  

; NAME/KEY: unsure  

; LOCATION: (1)..(260)  

; OTHER INFORMATION: unsure at all Xaa locations  

; FEATURE:  

; OTHER INFORMATION: Clone ID: MET4577_170575C.1.pep  

US-10-425-115-262020  
  

Query Match          30.7%; Score 48.5; DB 17; Length 260;  

Best Local Similarity 35.5%; Pred No. 59;  

Matches   11; Conservative    7; Mismatches 10; Indels   3; Gaps   1;  
  

QY       5 KGFDLITNFQQVADAL---NISLLPNPLATA 32  

        :|::|::|::|::|::|::|::|::|::|  

DB      179 RGIDPVTRPIAADAVTTVTVSFQPSPSAAA 209  
  

RESULT 12  

US-10-263-929-133  

; Sequence 133, Application US/10263929  

; Publication No. US20040067535A1  

; GENERAL INFORMATION:  

; APPLICANT: Kim, Jaeseob  

; APPLICANT: Galant, Ron  

; TITLE OF INVENTION: Alzheimer's Disease Linked Genes  

; FILE REFERENCE: LSD-07417  

; CURRENT APPLICATION NUMBER: US/10/263,929  

; CURRENT FILING DATE: 2002-10-03  

; NUMBER OF SEQ ID NOS: 213  

; SOFTWARE: PatentIn version 3.2  

; SEQ ID NO 133  

; LENGTH: 1800  

; TYPE: PRT  

; ORGANISM: Homo sapiens  

US-10-263-929-133  
  

Query Match          30.7%; Score 48.5; DB 15; Length 1800;  

Best Local Similarity 34.5%; Pred No. 6,5e+02;  

Matches   10; Conservative    7; Mismatches 11; Indels   1; Gaps   1;  
  

QY       2 DTIKGFDLITNPFQVVADALNISLLPNPLA 30  

        :|::|::|::|::|::|::|::|::|  

DB     1141 EQLGSHLELNHQ-IADSMEFGFLPNPVA 1168  
  

RESULT 13  

US-10-267-502-242  

; Sequence 242, Application US/10267502  

; Publication No. US20040071700A1  

; GENERAL INFORMATION:
```



```
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 172368
/ LENGTH: 231
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(231)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_70510C.1.pep
US-10-437-963-172368
```

```
Query Match      30.4%; Score 48; DB 16; Length 231;
Best Local Similarity 48.3%; Pred.No. 61;
Matches 14; Conservative 2; Mismatches 11; Indels 2; Gaps 1;
```

```
QY      6 GFDLITNFQVVADALNISLLPN--PLATA 32
      ||||| ||| : ||| : |||
Db     15 GFDLTKNFQDNDSDAFFRSVKERVVPLQKA 43
```

```
Search completed: November 11, 2004, 07:41:45
Job time : 77.3238 secs
```

This Page Blank (uspto)

RESULT 7
AH1220
acetate kinase homolog AckA2 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 16-Aug-2004
C:Accession: AH1220
R:Glaser, P.; Franquel, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kraft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vasquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1220
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <GLA>
A:Cross-references: UNIPROT:Q877V1; GB:NC_003210; PIDN:CAC99246.1; PID:gi6410584; GSPDB:AB01
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: AckA2
C:Superfamily: Acetate/propionate kinase
Query Match 30.7%; Score 48.5; DB 2; Length 397;
Best Local Similarity 28.9%; Pred. No. 29;
Matches 11; Conservative 8; Mismatches 12; Indels 7; Gaps 1;
Qy 2 DTIKGFDLITNFQVADALNI-----SLPNPLATA 32
Db 106 EVKGIQSAVTNLAHPNANIIGIKTPRELLPNAVSVA 143
RESULT 8
S02138
arginine deiminase (EC 3.5.3.6) - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C:Accession: S02138; B82999
R:Baur, H.; Luehti, E.; Stalton, V.; Mercenier, A.; Haas, D.
Eur. J. Biochem. 179, 53-60, 1989
A:Title: Sequence analysis and expression of the arginine-deiminase and carbamate-kinase
A:Reference number: S02137; MUID:89137094; PMID:2537202
A:Accession: S02138
A:Molecule type: DNA
A:Residues: 1-418 <BAU>
A:Cross-references: UNIPROT:P13981; EMBL:X14694; NID:945285; PIDN:CAA32824.1; PID:945286
A:Note: part of this sequence, including the amino end of the mature protein, was confirmed by Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen.
A:Reference number: B82950; MUID:20437337; PMID:10984043
A:Accession: B82999
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-418 <STO>
A:Cross-references: GB:AB004930; GB:AB004091; NID:99951472; PIDN:AAG08556.1; GSPDB:GN001
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: arcA; PA5171
C:Superfamily: arginine deiminase
C:Keywords: hydrolase
F:2-418/Product: arginine deiminase #status predicted <MAT>
Query Match 30.4%; Score 48; DB 2; Length 418;
Best Local Similarity 40.6%; Pred. No. 36;
Matches 13; Conservative 7; Mismatches 8; Indels 4; Gaps 2;
Qy 1 MDTKGF---DLITNF-QVADALNISLEPNP 28
Db 279 LDTVFSFCDRLVTVFPEVKVEIVFSLRPDPF 310

RESULT 9
VCVWFS
env polyprotein - mink cell focus-forming virus
N:Alternate names: coat polyprotein
N:Contains: knob protein gp70; R protein; spike protein p15E
C:Species: mink cell focus-forming virus
C>Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 23-Aug-1997
C:Accession: A31668
R:Chattopadhyay, S.K.; Baroudy, B.M.; Holmes, K.L.; Fredrickson, T.N.; Lander, M.R.; Moyle, J.; Title: Biologic and molecular genetic characteristics of a unique MCF virus that is highly infectious.
A:Reference number: A31668; MUID:89085614; PMID:2535909
A:Accession: A31668
A:Molecule type: DNA
A:Residues: 1-636 <CHA>
C:Genetics:
A:Gene: env
C:Superfamily: type C retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; spike protein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-440/Product: knob protein gp70 #status predicted <GP7>
F:441-620/Product: spike protein p15E #status predicted <P1E>
F:621-636/Product: R protein #status predicted <RPT>
F:43,58,297,329,369,405/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 30.4%; Score 48; DB 1; Length 636;
Best Local Similarity 36.7%; Pred. No. 60;
Matches 11; Conservative 4; Mismatches 11; Indels 4; Gaps 1;
Qy 6 GFDLITNFQVADALNIS-----LLPNPLAT 31
Db 212 GIDPVTFSLTRQVLNIGFRLPIGNFVIT 241
RESULT 10
S13580
collagen alpha 1(IX) chain precursor, long splice form - human
N:Alternate names: procollagen alpha 1(IX) chain, long splice form
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C:Accession: S13580; S02140; C35980; S21087; S74294
R:Muragaki, Y.; Kimura, T.; Ninomiya, Y.; Olsen, B.R.
Eur. J. Biochem. 192, 703-708, 1990
A:Title: The complete primary structure of two distinct forms of human alpha-1(IX) collagen.
A:Reference number: S13580; MUID:91006164; PMID:2209617
A:Accession: S13580
A:Molecule type: mRNA
A:Residues: 1-931 <MR>
A:Cross-references: UNIPROT:P20849; EMBL:X54412; NID:G30083; PIDN:CAA38276.1; PID:G3008
R:Ninomiya, Y.; Castagnola, P.; Gerecke, D.; Gordon, M.K.; Jacenko, O.; LuValle, P.; McManus, B.T.; Olsen, B.R.
in Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp.79-114, Academic Pr
A:Title: The molecular biology of collagens with short triple-helical domains.
A:Reference number: S22243
A:Accession: S23295
A:Status: not compared with conceptual translation
A:Residues: 1-253, 'v', 255-815;835-884 <NIN>
R:Kimura, T.; Mattel, M.G.; Stevens, J.W.; Goldring, M.B.; Ninomiya, Y.; Olsen, B.R.
Eur. J. Biochem. 179, 71-78, 1989
A:Title: Molecular cloning of rat and human type IX collagen cDNA and localization of t
A:Reference number: S02140; MUID:89137096; PMID:2465149
A:Accession: S02140
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 580-596;598-620, 'R', 622-813;835-884 <KIM>
R:Muragaki, Y.; Nishimura, I.; Henney, A.; Ninomiya, Y.; Olsen, B.R.
Proc. Natl. Acad. Sci. U.S.A. 87, 2400-2404, 1990
A:Title: The alpha(IX) collagen gene gives rise to two different transcripts in both m
A:Reference number: A35980; MUID:90207204; PMID:1690886
A:Accession: C35980

A:Molecule type: DNA
 A:Residues: 1-4:233-267 <MUR2>
 A:Cross-references: EMBL:M32135
 R:Olsen, B.R.
 Submitted to the EMBL Data Library, February 1990
 A:Reference number: S21087
 A:Accession: S21087
 A:Molecule type: DNA
 A:Residues: 1-4:233-248, 'T', 250-267 <OLS>
 A:Cross-references: EMBL:M32135
 R:Diab, M.; Wu, J.J.; Eyre, D.R.
 Biochem. J. 314, 327-332, 1996
 A:Title: Collagen type IX from human cartilage: a structural profile of intermolecular cross-links
 A:Reference number: S64673; MUID:96195147; PMID:8660302
 A:Accession: S74294
 A:Molecule type: protein
 A:Residues: 405-417 <DIA>
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit are subsequently O-glycosylated.
 C:Genetics:
 A:Gene: GDB:COL9A1
 A:Cross-references: GDB:119794; OMIM:120210
 A:Map position: 6q12-6q14
 A:Introns: 5/2; 232/3; 260/3; 267/3
 C:Complex: type IX collagen may be a heterotrimer of one alpha 1(IX) chain, one alpha 2(IX) chain, and one alpha 3(IX) chain.
 C:Function:
 A:Description: structural component of extracellular fibrous polymer associated with type I collagen
 A:Note: in chondrocytes the long splice form is predominantly produced
 C:Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; heterotrimer
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-931/Product: collagen alpha 1(IX) chain, long splice form #status predicted <MAT>
 F:24-268/Domain: non-collagenous NC4 #status predicted <NC4>
 F:269-405/Domain: non-collagenous NC3 #status predicted <NC3>
 F:406-417/Domain: non-collagenous NC2 #status predicted <NC2>
 F:418-756/Domain: non-collagenous NC2 #status predicted <NC2>
 F:757-786/Domain: non-collagenous NC2 #status predicted <NC2>
 F:787-901/Domain: non-collagenous COL1 #status predicted <COL1>
 F:902-931/Domain: non-collagenous NC1 #status predicted <NC1>
 F:171/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 30.4%; Score 48; DB 2; Length 931;
 Best Local Similarity 64.3%; Pred. No. 94;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 DTIKGFDLITNFQV 15
 Db 52 DDLPGFDLISQFQV 65

RESULT 11

S55123
 hypothetical protein YMR176w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YMR1010.06
 C:Species: Saccharomyces cerevisiae
 C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
 A:Accession: S55123
 R:Churcher, C.M.
 Submitted to the EMBL Data Library, June 1995
 A:Reference number: S55118
 A:Cross-references: UNIPROT:Q03214; EMBL:Z49808; NID:9854440; PIDN:CAA99909.1; PID:g8544
 A:Molecule type: DNA
 A:Residues: 1-1411 <CHU>
 A:Cross-references: UNIPROT:Q03214; EMBL:Z49808; NID:9854440; PIDN:CAA99909.1; PID:g8544
 A:Experimental source: strain AB972
 C:Genetics:
 A:Gene: SGD:ECM5
 A:Cross-references: SGD:S0004788; MIPS:YMR176w
 A:Map position: 13R

Query Match 30.4%; Score 48; DB 2; Length 1411;
 Best Local Similarity 33.3%; Pred. No. 1.5e+02;
 Matches 10; Conservative 11; Mismatches 7; Indels 2; Gaps 1;

QY 1 MDTIKG--FDLITNFQVADALNISLLPNP 28
 Db 1294 VDNIEGVTPELDLKRILVESLKLISLPDP 1323

RESULT 12

AC1204
 phosphotransferase system (PTS) beta-glucoside-specific enzyme IABC homolog lmo1035 [in
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 A:Accession: AC1204
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
 A:Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AC1204
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-618 <GLA>
 A:Cross-references: UNIPROT:Q8Y882; GB:NC_003210; PIDN:CAC99113.1; PID:g16410437; GSPDB
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo1035
 C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase

Query Match 30.1%; Score 47.5; DB 2; Length 618;
 Best Local Similarity 35.5%; Pred. No. 69;
 Matches 11; Conservative 8; Mismatches 5; Indels 7; Gaps 2;

QY 1 MDTIK--GPD-----LITNFQVADALNISL 24
 Db 573 LDTIKRAGYDITPIIVTNSATLADVIIVNL 603

RESULT 13

D90988
 hypothetical protein ECS2876 [imported] - Escherichia coli (strain O157:H7, substrain R
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 16-Aug-2004
 A:Accession: D90988
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
 A:Reference number: A99629; MUID:21156231; PMID:11238796
 A:Accession: D90988
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-791 <HAY>
 A:Cross-references: UNIPROT:Q8X397; GB:BA000007; PIDN:BA036299.1; PID:gl332345; GSPDB:
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: ECS2876
 C:Superfamily: Signal transduction protein with an integral membrane domain and Pas, GG

Query Match 30.1%; Score 47.5; DB 2; Length 791;
 Best Local Similarity 48.3%; Pred. No. 92;
 Matches 14; Conservative 4; Mismatches 8; Indels 3; Gaps 2;

QY 1 MDTIKGF--DLITNFQVADALNISLLPN 27
 Db 757 MDTLSIGIGVDLIYG-EVIADAPLDLLMN 784

RESULT 14

F88333
 partial probable sensor kinase Z3236 [imported] - Escherichia coli (strain O157:H7, sub
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Aug-2004

C;Accession: P85833
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: P85833
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-935 <STO>
 A;Cross-references: UNIPROT:Q8X7K1; GB:AB005174; NID:g12516270; PIDN:AAG57130.1; GSPDB:G
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:
 A;Gene: Z3236
 C;Superfamily: Signal transduction protein with an integral membrane domain and Pas, GDB

Query Match 30.1%; Score 47.5; DB 2; Length 935;
 Best Local Similarity 48.3%; Pred. No. 1.1e+02;
 Matches 14; Conservative 4; Mismatches 8; Indels 3; Gaps 2;

Qy 1 MDTIKGF--DLITNFQVVADALNISLLPN 27
 Db 901 MDTLSGIGVDLIYG-EVIADAPLDLLMN 928

RESULT 15
 B64973
 Yege protein - Escherichia coli (strain K-12)
 C;Species: Escherichia coli
 C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 16-Aug-2004
 C;Accession: B64973
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503
 A;Accession: B64973
 A;Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-1105 <BLAT>
 A;Cross-references: UNIPROT:P38097; GB:AB000296; GB:U00096; NID:g1788373; PIDN:AAC75128.
 A;Experimental source: strain K-12, substrain MG1655
 C;Genetics:
 A;Gene: yege
 C;Superfamily: Signal transduction protein with an integral membrane domain and Pas, GDB

Query Match 30.1%; Score 47.5; DB 2; Length 1105;
 Best Local Similarity 48.3%; Pred. No. 1.4e+02;
 Matches 14; Conservative 4; Mismatches 8; Indels 3; Gaps 2;

Qy 1 MDTIKGF--DLITNFQVVADALNISLLPN 27
 Db 1071 MDTLSGIGVDLIYG-EVIADAPLDLLVN 1098

Search completed: November 10, 2004, 15:55:05
 Job time : 11.9324 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 10, 2004, 14:50:40 ; Search time 53.637 Seconds
(without alignments)
343.270 Million cell updates/sec

Title: US-10-092-750-67

Perfect score: 158

Sequence: 1 MDTIKGFDLITNFQVADALNISLLPNPLATA 32

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57.5	36.4	662	Q6CON9	O6CON9 yarrowia li
2	53	33.5	309	Q6L2H0	Q6L2H0 picophilus
3	52	32.9	321	Q9YD78	Q9YD78 aeropyrum p
4	52	32.9	377	Q6D3R8	Q6D3R8 erwinia car
5	52	32.9	487	Q6LQ81	Q6LQ81 photobacter
6	52	32.9	487	CAG20545	CAG20545 photobact
7	52	32.9	636	Q85506	Q85506 murine leuk
8	52	32.9	707	Q6N3Y2	Q6N3Y2 corynebacte
9	52	32.9	707	CAB48764	CAB48764 corynebac
10	51	32.3	209	Q74NJ5	Q74NJ5 nanaoarchaeu
11	51	32.3	209	AAR39201	AAR39201 nanaoarcha
12	50.5	32.0	268	O80213	O80213 methanobact
13	50.5	32.0	323	O36002	O36002 monocetomo
14	50.5	32.0	351	Q8KCF2	Q8KCF2 chlorobium
15	50.5	32.0	519	Q6FCV0	Q6FCV0 acinetobact
16	50.5	32.0	534	Q6FA70	Q6FA70 acinetobact
17	50.5	32.0	624	Q9YV11	Q9YV11 melanoplus
18	50	31.6	314	Q6N1X8	Q6N1X8 corynebacte
19	50	31.6	314	CAE49155	CAE49155 corynebact
20	50	31.6	629	Q7RH66	Q7RH66 plasmodium
21	50	31.6	774	Q6ER20	Q6ER20 candida gla
22	50	31.6	1054	Q6BTD4	Q6BTD4 debaryomyce
23	49	31.0	113	Q9RH08	Q9RH08 streptomyce
24	49	31.0	141	Q71S22	Q71S22 scombridae
25	49	31.0	141	AAQ14279	AAQ14279 scombridae
26	49	31.0	294	O50152	O50152 streptomyce
27	49	31.0	376	Q89622	Q89622 clostridium
28	49	31.0	481	PRTB_ERWCH	P16316 erwinia chr
29	49	31.0	574	Q7RMX3	Q7RMX3 plasmodium
30	49	31.0	644	Q6FIT9	Q6FIT9 mesoplasma
31	48.5	30.7	120	Q88914	Q88914 pseudomonas

32 48.5 30.7 255 1 MYB3 MAIZE
33 48.5 30.7 324 2 O15556
34 48.5 30.7 350 2 Q27820
35 48.5 30.7 355 2 Q7XYU3
36 48.5 30.7 397 1 ACK2 LISMO
37 48.5 30.7 397 2 Q720R2
38 48.5 30.7 397 2 AAT03952
39 48.5 30.7 406 2 Q894W1
40 48.5 30.7 893 2 Q9SXH2
41 48.5 30.7 1729 1 NU98 HUMAN
42 48 30.4 109 2 Q7MTF3
43 48 30.4 113 2 Q75UW4
44 48 30.4 113 2 BAD08678
45 48 30.4 221 2 Q8ID93

ALIGNMENTS

RESULT 1

Q6CON9 PRELIMINARY; PRT; 662 AA.
AC Q6CON9; (TREMELrel. 28, Created)
DT 01-OCT-2004 (TREMELrel. 28, Last sequence update)
DT 01-OCT-2004 (TREMELrel. 28, Last annotation update)
DE Similar to sp|P08540 Kluyveromyces fragilis Potential acid
DE phosphate.
GN ORFNames=YALIOF23067g;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Matck C., Neuveglise C., Talia E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boiarane A., Boyer J., Cattolico L., Confaniolero F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet F., Kachouri R.,
RA Kerret A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swenne D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarfelli C., Gaillardin C., Weissenbach J.,
Wincker P., Souciet J.L.;
RL "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382132; CAG79584.1; -
SQ SEQUENCE 662 AA; 71355 MW; 75CA81C084A5F96B CRC64;

Query Match 36.4%; Score 57.5; DB 2; Length 662;

Best Local Similarity 46.2%; Pred No. 16;
Matches 12; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

QY 1 MDTIKGFDLITNFQVADALNISLL 25

305 LDHILGRFDTVANPFLKVLADSLNISIV 330

RESULT 2

Q6L2H0 PRELIMINARY; PRT; 309 AA.
ID Q6L2H0


```

Db      277 VQGFSLQQGFAQYVDGFNISM L 298
::||| | || | |||:|
RESULT 6
CAG20545 PRELIMINARY; PRT; 487 AA.
ID CAG20545 AC CAG20545;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DE Putative Na+/H+ antiporter.
GN MK2219 OR PPRA2147..
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ss9;
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
RA Cestaro A., Malacrida G., Simionati B., Cannata N., Bartlett D.,
RA Valle G.;
RT "Genome Analysis of Photobacterium profundum reveals the complexity of
RT high pressure adaptations.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ss9;
RA Cestaro A.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR ENBL; CR378670; CAG20545.1; -.
SQ SEQUENCE 487 AA; 51332 MW; D56453B148B800DA CRC64;

Query Match 32.9%; Score 52; DB 2; Length 487;
Best Local Similarity 45.5%; Pred.No. 76;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 4 IKGPDLITNFQVVADALNISLL 25
::||| | |||:|
Db      277 VQGFSLQQGFAQYVDGFNISM L 298
::||| | || | |||:|

RESULT 7
Q85506 Q85506 PRELIMINARY; PRT; 636 AA.
ID ID Q85506 AC Q85508;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Env polyprotein.
OS Murine leukemia virus.
OC Viruses; Retroid viruses; Retroviridae; Gammaetrovirus.
OX NCBI_TaxID=11786;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83059868; Pubmed=7143566;
RA Bosselman R.A., van Straaten F., van Beveren C.P., Verma I.M.,
RA Vogt M.;
RT "Analysis of the env gene of a molecularly cloned and biologically
RT active Moloney mink cell focus-forming proviral DNA.";
RL J Virol 44:19-31(1982).
DR HBSPL; J02254; AAA46517.1; -.
DR HSP; P03385; IMOF.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPRO02050; Env_polyprotein.
DR InterPro; IPRO08981; FMWVrecept-bind.
DR Pfam; PF00429; TLY_coat; 1.
FT CHAIN 1 440 Potential.
SQ SEQUENCE 636 AA; 69461 MW; FB24248A547CA81 CRC64;

```

```

RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA De Zoyea A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitz E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.;
RA "The complete genome sequence and analysis of Corynebacterium
RT diptheriae NC7C13129.";
RL Nucleic Acids Res. 31:6516-6523(2003).
RW EMBL; BX248354; CAE48764.1; -.
KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 707 AA; 75874 MW; 9AD805227F815E46 CRC64;

Query Match 32.9%; Score 52; DB 2; Length 707;
Best Local Similarity 41.9%; Pred.No. 1.le+02;
Matches 13; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 MDTIKGFDLITNFQVVDALNLSILNPPLAT 31
Db 253 IDLAGNDNAALFQVDDATEAGLDPERPAT 283
: : | | | | | : : | |
PRELIMINARY; PRT; 209 AA.

RESULT 10
Q74NJ5 PRELIMINARY; PRT; 209 AA.
ID Q74NJ5 AC Q74NJ5;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE NEQ352.
GN OrderedLocusNames=NEQ352;
OS Nanoarchaeum equitans.
OC Archaea; Nanoarchaeota; Nanoarchaeum.
OX NCBI_TaxID=160232;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kin4-M;
RX MEDLINE=22946215; PubMed=14560662;
RA Waters E., Hohn M.J., Ahel I., Graham D.E., Adams M.D., Barnstead M.,
RA Beeson K.Y., Bibbs L., Bolanos R., Keller M., Kretz K., Lin X.,
RA Mathur E., Ni J., Podar M., Richardson T., Sutton G.G., Simon M.,
RA Soell D., Stetter K.O., Short J.M., Noorderwier M.;
RT "The genome of Nanoarchaeum equitans: insights into early archaeal
RT evolution and derived parasitism.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:12984-12988(2003).
DR EMBL; AE017199; AAR39201.1; -.
DR InterPro; IPR002749; DUF63.
DR Pfam; PF01889; DUF63; 1.
KW Complete proteome.
SQ SEQUENCE 209 AA; 23996 MW; 650B36C348A59967 CRC64;

Query Match 32.3%; Score 51; DB 2; Length 209;
Best Local Similarity 40.0%; Pred.No. 45;
Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 2 DTIKGFDLITNFQVVDALNLSILNPPLAT 31
Db 32 ETTKFFLLSLFIRVFVDINLPNTFTT 61
: : | | | | | : : | | | | |
PRELIMINARY; PRT; 209 AA.

RESULT 11
AAR39201 PRELIMINARY; PRT; 209 AA.
ID AAR39201 AC AAR39201;
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE NEQ352.
GN NEQ352.
OS Nanoarchaeum equitans.
OC Archaea; Nanoarchaeota; Nanoarchaeum.
OX NCBI_TaxID=160232;
RN [1]
RP SEQUENCE FROM N.A.
```

J. Mol. Evol. 47:190-199(1998).

-I- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate + NAD(+)= 3-phospho-D-glyceroyl phosphate + NADH.

-II- PATHWAY: Second phase of glycolysis; first step.

-III- SUBUNIT: Homotetramer (By similarity).

-IV- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

-V- SIMILARITY: Belongs to the glyceraldehyde-3-phosphate dehydrogenase family.

EMBL; AF022420; AAC63603.1; -. HSSP; P17721; IHGD.

GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p...); IEA.

GO; GO:0016491; F:oxygenoreductase activity; IEA.

GO; GO:0006096; P:glycolysis; IEA.

InterPro; IPR000173; GAP_dhrogenase.Fam; PF02800; Gp_dh_C_1.Pfam; PF00044; Gp_dh_N_1.DR PRINTS; PR00078; G3PDHDRGNASE.DR PROSITE; PS00071; GAPDH; 1.KW Glycolysis; NAD; Oxidoreductase.EF NON_TER 1 323 FT NON_TER 323 323

SEQUENCE 323 AA; 3F6DBFL7FE5C6792 CRC64;

Query Match 32.0%; Score 50.5; DB 2; Length 323;
Best Local Similarity 41.4%; Pred.No.83;
Matches 12; Conservative 4; Mismatches 12; Indels 1; Gaps 1;

QY 1 MDRIKGFDLI-TNFOVVADALNISLLPNP 28
||| ||| : | : : :
D5 285 MDPIVSDDIDDNHSSIVDGLSTWLVPNP 313

RESULT 14

Q8KCZF PRELIMINARY; PRT; 351 AA.

ID Q8KCZF

AC Q8KCZF2

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Rod shape-determining protein MreB.

GN NameMreB-2; OrderedLocusNames=CTL1470;

OS Chlorobium tepidum.

OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;

OC Cholorobaculum.

OX NCBI_TaxId=1097;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=TLS / ATCC 49652 / DSM 12025.

RX MEDLINE=22103685; PubMed=1203901; DOI=10.1073/pnas.132181499;

RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,

RA Dodson R.J., DeBoy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,

RA Hickney E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,

RA Holt I.E., Umavaj L.A., Mason T.M., Brenner M., Shea T.P.,

RA Parkey D.S., Niemman W.C., Feldblum T.V., Hansen C.L., Craven M.B.,

RA Radune D., Vamathevan J.J., Khouri H.M., White O., Gruber T.M.,

RA Kechum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.N.;

RT "The complete genome sequence of Chlorobium tepidum TLS, a

RT photosynthetic, anaerobic, green-sulfur bacterium."

RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).

DR EMBL; AE012904; AAM72697.1; -. DR EMBL; CT1470; -.

DR InterPro; IPR004753; MreB_Mrl.

DR Pfam; PF06723; MreB_Mbl; I.

DR PRINTS; PR01652; SHAPEPROTEIN.

DW Complete proteome.

XW SEQUENCE 351 AA; 379973 MW; 7A7211FF855C625E CRC64:

Query Match 32.0%; Score 50.5; DB 2; Length 351;
Best Local Similarity 36.7%; Pred. No. 90;
Matches 11; Conservative 8; Mismatches 8; Indels 3; Gaps 1;

```

Db      232   TVKGFNLLTGFPDQTGISTAVLRRIITTPPL 261

RESULT 15
Q6FCV0
ID       Q6FCV0          PRELIMINARY;           PRT;        519 AA.
AC       Q6FCV0;         TEMBLrel. 27, Created)
DT     05-JUL-2004    (TEMBLrel. 27, Last sequence update)
DT     05-JUL-2004    (TEMBLrel. 27, Last annotation update)
DE Alkyl hydroperoxide reductase subunit F [BC 1.8.1.-).
GN Name=ahpf; OrderedLocusNames=ACIAD1234;
DB Acinetobacter sp. (strain ADP1).
OS Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
[1]
SEQUENCE FROM N.A.
RA Barbe V., Vallent D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre I., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Christon L.N., Weissenbach J., Marlier P., Cohen G.N., Medigue C.;
RT "Unique features revealed by the genome sequence of Acinetobacter sp."
RT ADP1, a versatile and naturally transformation competent bacterium.",
RL Nucleic Acids Res. 0:0-0(2004).
DR ENML; CR543861; CAG68109.1; -.
DR InterPro; IPR000759; Adrmrx_reductase.
DR InterPro; IPR001327; PAD_pyr_redox.
DR InterPro; IPR00103; Pyridine_redox_2.
DR InterPro; IPR001100; Pyr_redox.
DR InterPro; IPR008255; Pyr_redox2_AS.
DR InterPro; IPR003042; Rng_mnoxxygenase.
DR Pfam; PF00070; Pyr_redox; 1.
DR PRINTS; PR00419; ADXRDTASE.
DR PRINTS; PR00368; PADPNR.
DR PRINTS; PR00411; PNDRTASEII.
DR PRINTS; PR00469; PNDRTASEII.
DR PRINTS; PR00420; RNGMNOXGNASE.
DR PROSITE; PS00573; PYRIDINE_REDOX_2; 1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 519 AA; 55803 MW; 0CA26DB974FA8F18 CRC64;

Query Match              32.0%; Score 50.5; DB 2; Length 519;
Best local similarity    32.6%; Pred.No.1.4e+02;
Matches 14; Conservative Mismatches 11; Indels 11; Gaps

Qy      1 MDTIKGFDLTINPOV-----VADALNISILPNPLATA 32
            | : | | | | | | | | | | | | | | | |
Db      108 LNIGKLNLIKADPDFVSLSCHNCPPDVQAALNIAYNFNSTA 150
            | : | | | | | | | | | | | | | | | |

Search completed: November 10, 2004, 15:53:28
Job time : 55.7139 secs
```

Qy

1 MDTIKGFDLITNFQV-----VADALNISLLPNPLATA 32
:: ||| : | : |
||| : | : ||| : | :

pB

108 LNOIKGLNLKADFDFVSLSCHNCPDVAALNDLIAYNPNSTA 150

Search completed: November 10, 2004, 15:53:28
Job time : 55.7139 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM.protein - protein search, using sw model

Run on: November 10, 2004, 14:55:47 ; Search time 10.1779 Seconds
(without alignments)
143.349 Million cell updates/sec

Title: US-10-092-750-68
Perfect score: 118
Sequence: 1 ATWKTQLGLLDRIQAPFSSPH 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	42.4	339	4	US-09-148-545-246
2	49.5	41.9	352	4	US-09-134-000C-5243
3	47	39.8	369	4	US-09-107-532A-5270
4	46.5	39.4	362	4	US-09-107-532A-7093
5	46	39.0	732	3	US-08-989-299-5
6	46	39.0	732	4	US-09-407-427-5
7	46	39.0	1312	4	US-08-989-299-8
8	46	39.0	1312	4	US-09-407-427-8
9	45	38.1	349	4	US-09-252-991A-27544
10	45	38.1	1313	3	US-08-989-299-9
11	45	38.1	1313	4	US-09-407-427-9
12	45	38.1	3433	4	US-09-091-501B-10
13	45	38.1	3433	4	US-09-538-092-1136
14	44.5	37.7	153	4	US-09-621-976-4517
15	44	37.3	100	3	US-09-227-357-617
16	44	37.3	284	4	US-09-489-039A-8653
17	44	37.3	732	1	US-08-481-826-2
18	44	37.3	732	3	US-08-989-299-4
19	44	37.3	732	4	US-09-407-427-4
20	44	37.3	751	3	US-09-036-987A-24
21	44	37.3	751	3	US-09-370-700-24
22	44	37.3	751	4	US-09-603-207-24
23	44	37.3	1306	3	US-08-989-299-7
24	44	37.3	1306	4	US-09-407-427-7
25	43.5	36.9	144	4	US-09-513-999C-5382
26	43	36.4	175	3	US-08-817-441-98
27	43	36.4	358	1	US-08-604-913B-11

28	43	36.4	521	1	US-08-276-213-3
29	43	36.4	765	4	US-09-489-039A-5204
30	42	35.6	283	4	US-09-252-991A-32069
31	42	35.6	700	4	US-08-933-711B-5
32	41	34.7	82	4	US-09-248-796A-19112
33	41	34.7	88	4	US-09-621-976-4578
34	41	34.7	88	4	US-09-621-976-4579
35	41	34.7	89	4	US-09-621-976-6841
36	41	34.7	159	4	US-09-328-352-5862
37	41	34.7	162	4	US-09-583-110-4507
38	41	34.7	382	4	US-09-543-681A-5257
39	41	34.7	445	4	US-09-489-039A-12278
40	41	34.7	445	4	US-09-710-279-1080
41	41	34.7	591	3	US-09-134-001C-3275
42	41	34.7	754	4	US-09-252-991A-33133
43	40.5	34.3	503	4	US-09-107-532A-3708
44	40.5	34.3	3080	6	5223423-4
45	40	33.9	88	4	US-09-270-767-34558

ALIGNMENTS

RESULT 1
US-09-148-545-246
; Sequence 246, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rozen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23

/	EARLIER APPLICATION NUMBER:	60/047,584
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,500
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,587
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,492
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,598
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,613
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,582
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/047,596
/	EARLIER FILING DATE:	1997-05-23
/	EARLIER APPLICATION NUMBER:	60/043,580
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,568
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,314
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,569
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,311
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,671
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,674
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,669
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,312
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,313
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,672
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/043,315
/	EARLIER FILING DATE:	1997-04-11
/	EARLIER APPLICATION NUMBER:	60/048,974
/	EARLIER FILING DATE:	1997-06-06
/	EARLIER APPLICATION NUMBER:	60/056,886
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,877
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,889
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,893
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,630
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,878
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,662
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,872
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,882
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,637
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,903
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,888
/	EARLIER FILING DATE:	1997-08-22
/	EARLIER APPLICATION NUMBER:	60/056,879

; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 246
; LENGTH: 339

Query Match 42.4%; Score 50; DB 4; Length 339;
Best Local Similarity 50.0%; Pred. No. 3.1;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TWKMTLQGLLDRIQAF 17
||:||||:|
Db 223 TWQALTGALLERMQTY 238

RESULT 2

US-09-134-000C-5243
; Sequence 5243, Application US/09134000C
; Patent No. 6617156

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: Patent in version 3.1
SEQ ID NO 5243
LENGTH: 352
TYPE: PRT
ORGANISM: Enterococcus faecalis

US-09-134-000C-5243

Query Match 41.9%; Score 49.5; DB 4; Length 352;
Best Local Similarity 40.0%; Pred. No. 3.9;
Matches 10; Conservative 4; Mismatches 6; Indels 5; Gaps 1;

Qy 2 TW-----MKTQLGLLDRIQAPSSP 21
||:||||:|
Db 253 TWGLGSMPTWEGTEKLYAIPGSP 277

RESULT 3

US-09-107-532A-5270
; Sequence 5270, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5270:

SEQUENCE CHARACTERISTICS:

LENGTH: 369 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...369

SEQUENCE DESCRIPTION: SEQ ID NO: 5270:

US-09-107-532A-5270

Query Match 39.8%; Score 47; DB 4; Length 369;

Best Local Similarity 42.9%; Pred. No. 11;

Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 2 TWKMTLQGLLDRIQAPSSPH 22

Db 178 TWRTVQGISDRKTVFNPNH 198

RESULT 4

US-09-107-532A-7093

; Sequence 7093, Application US/09107532A

; Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 7093:

SEQUENCE CHARACTERISTICS:

LENGTH: 362 amino acids

TYPE: amino acid

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...362
; SEQUENCE DESCRIPTION: SEQ ID NO: 7093:
US-09-107-532A-7093

Query Match          39.4%; Score 46.5; DB 4; Length 362;
Best Local Similarity 44.0%; Pred. No. 13;
Matches 11; Conservative 1; Mismatches 8; Indels 5; Gaps 1;

QY  2 TW-----MKTQLGLLDRIQAPSSP 21
    |||||
Db   257 TWGLGSMPTLDSANDRLYAPGSP 281
    |||||

RESULT 5
US-08-989-299-5
; Sequence 5, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Artold E., Beth
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-025.01
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-989-299-5

Query Match          39.0%; Score 46; DB 3; Length 732;
Best Local Similarity 35.0%; Pred. No. 35;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY  3 WKKTLQGLLDRIQAPSSPH 22
    |||
Db   309 WAQTWSNIYDLVAPFSPAPN 328
    |||

RESULT 6
US-09-407-427-5
; Sequence 5, Application US/09407427
; Patent No. 6610497
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Robinson, Keith E.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND DIAGNOSTIC USES THEREFOR
; FILE REFERENCE: MNI-132CP2
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US/09/407,427
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 08/989,299
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Murine sp.
US-09-407-427-5

Query Match          39.0%; Score 46; DB 4; Length 732;
Best Local Similarity 35.0%; Pred. No. 35;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY  3 WKKTLQGLLDRIQAPSSPH 22
    |||
Db   309 WAQTWSNIYDLVAPFSPAPN 328
    |||

RESULT 7
US-08-989-299-8
; Sequence 8, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Robinson, Keith E.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Artold E., Beth
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-025.01
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1312 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-989-299-8
```

```

RESULT 10
US-08-989-299-9
; Sequence 9, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E., Beth
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-025.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1313 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-989-299-9

Query Match      38.1%; Score 45; DB 3; Length 1313;
Best Local Similarity 36.8%; Pred.No. 1e-02; 8; Indels 0; Gaps 0;
Matches 7; Conservative 4; Mismatches 8;

QY          3 WVKTLQLGLDRIQAFPSPP 21
           |||::|||::|||::
DB          890 WAQTWSNIYDLVAPFSAP 908

RESULT 11
US-09-407-427-9
; Sequence 9, Application US/09407427
; Patent No. 6610497
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG AND THE
; TITLE OF INVENTION: AND DIAGNOSTIC USES THEREFOR
; FILE REFERENCE: MMI-132CP2
; CURRENT APPLICATION NUMBER: US/09/407,427
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 09/163,648
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 08/989,299
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1313
; TYPE: PRT
; ORGANISM: Rattus sp.
```

```

US-09-407-427-9
Query Match      38.1%; Score 45; DB 4; Length 1313;
Best Local Similarity 36.8%; Pred. No. 1e+02;
Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY      3 WMKTLQGLLDRIQAPSSP 21
DB      890 WAQTSNIYDLVAPFSP 908

RESULT 12
US-09-091-501B-10
; Sequence 10, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathon M
; APPLICANT: Davies, Kay E
; TITLE OF INVENTION: Utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091,501B
; CURRENT FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3433
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (239) ... (250)
; OTHER INFORMATION: Description of Artificial Sequence: Full length
; OTHER INFORMATION: utrophin construct; Xaa = unknown
US-09-091-501B-10

Query Match      38.1%; Score 45; DB 4; Length 3433;
Best Local Similarity 40.0%; Pred. No. 3.1e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      1 ATWMKTLQGLLDRIQAPSS 20
DB      1665 STWLYQAEALLDEIEKKPTS 1684

RESULT 13
US-09-538-092-1136
; Sequence 1136, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1136
; LENGTH: 3433
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-538-092-1136

Query Match      38.1%; Score 45; DB 4; Length 3433;
Best Local Similarity 40.0%; Pred. No. 3.1e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY      1 ATWMKTLQGLLDRIQAPSS 20
DB      1665 STWLYQAEALLDEIEKKPTS 1684

RESULT 14
US-09-621-976-4517
; Sequence 4517, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4517
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-4517

Query Match      37.7%; Score 44.5; DB 4; Length 153;
Best Local Similarity 52.9%; Pred. No. 10;
Matches 9; Conservative 6; Mismatches 1; Indels 1; Gaps 1;

QY      4 MKTLQGLLDRIQAPSS 20
DB      112 LKTL-GVLEKIQAYPEA 127

RESULT 15
US-09-227-357-617
; Sequence 617, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916

```